

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 12, 2006, 06:25:40 ; Search time 19961 Seconds  
(without alignments)  
20089.779 Million cell updates/sec

Title: US-09-522-753-4  
Perfect score: 8564.8  
Sequence: 1 catgctgggctccacacagc.....caaaaaaaaaaaaaaa 8571

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_esc1:\*  
2: gb\_esc2:\*  
3: gb\_esc3:\*  
4: gb\_hnc:\*  
5: gb\_esc4:\*  
6: gb\_esc5:\*  
7: gb\_esc6:\*  
8: gb\_esc7:\*  
9: gb\_gsa1:\*  
10: gb\_gsa2:\*  
11: gb\_gsa3:\*

Pred. No. is the number of results predicted by have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5140.2	60.0	7372	10	AY412686 Homo sapi
2	4216.6	49.2	6836	10	AY412687 Pan trogl
3	3315.2	38.7	7013	10	AY412688 Mus muscu
4	1473.6	17.2	2000	4	BC033087 Homo sapi
5	1297.4	15.1	1362	4	BC020427 Homo sapi
6	887.8	10.4	1093	3	BM909096 AGENCOURT
7	868.6	10.1	965	3	BM555371 AGENCOURT
8	841.8	9.8	846	5	BU557144 AGENCOURT
9	815.6	9.5	1126	3	BM553310 AGENCOURT
10	798.2	9.3	984	5	BQ690869 AGENCOURT
11	796.4	9.3	1051	3	BM423558 AGENCOURT
12	796.2	9.3	1159	3	BM802749 AGENCOURT
13	795.2	9.3	1057	3	BM477568 AGENCOURT
14	787.6	9.2	1022	3	BM910785 AGENCOURT
15	782	9.1	875	5	BQ711119 AGENCOURT
16	779	8.9	1010	3	BM915686 AGENCOURT
17	766.2	8.9	1030	3	BM558844 AGENCOURT
18	760.8	8.9	882	5	BX397973 AGENCOURT
19	754.6	8.8	892	5	BU538827 AGENCOURT
20	748.8	8.7	1036	5	BM910476 AGENCOURT
21	747.6	8.7	868	5	BX368971 AGENCOURT
22	746.8	8.7	882	5	BU180236 AGENCOURT

23	746	8.7	923	5	BX390462
24	743.4	8.7	1075	3	BM461469 AGENCOURT
25	740	8.6	916	5	BQ892847 AGENCOURT
26	737.4	8.6	1045	3	BM560255 AGENCOURT
27	730.8	8.5	1038	3	BM910704 AGENCOURT
28	730.6	8.5	1066	3	BM471347 AGENCOURT
29	728.4	8.5	882	2	BI089430 AGENCOURT
30	720.8	8.4	769	5	BX368972 AGENCOURT
31	712.6	8.3	988	5	BU172348 AGENCOURT
32	711.8	8.3	958	2	BG831424 AGENCOURT
33	711.4	8.3	880	8	CV806970 AGENCOURT
34	711	8.3	875	5	BQ691710 AGENCOURT
35	707.4	8.3	923	5	BU184403 AGENCOURT
36	706.8	8.3	1094	3	BM560912 AGENCOURT
37	703	8.2	876	2	BG252161 AGENCOURT
38	703	8.2	1067	3	BM472005 AGENCOURT
39	700.6	8.2	928	5	BU164114 AGENCOURT
40	695.8	8.1	772	6	CF135847 AGENCOURT
41	692.8	8.1	727	5	BU632778 AGENCOURT
42	692	8.1	830	7	CO648233 AGENCOURT
43	688	8.0	861	5	BU191146 AGENCOURT
44	683	8.0	737	6	CD742837 AGENCOURT
45	679.6	7.9	1087	2	BG252257 AGENCOURT

#### ALIGNMENTS

RESULT 1  
AY412686  
LOCUS  
DEFINITION  
Homo sapiens NCOR2 gene, VIRTUAL TRANSCRIPT, partial sequence.  
ACCESSION  
AY412686.1 GI:39768651  
VERSION  
GSS.  
KEYWORDS  
Homo sapiens (human)  
SOURCE  
Homo sapiens  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 7372)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,  
Fieritz,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
Adams,M.D. and Cargill,M.  
Interfering nonneutral evolution from human-chimp-mouse orthologous  
gene ctios  
JOURNAL  
Science 302 (5652), 1960-1963 (2003)  
PUBMED  
14671302  
REFERENCE  
2 (bases 1 to 7372)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,  
Fieritz,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
Adams,M.D. and Cargill,M.  
Direct Submision  
JOURNAL  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source  
1..7372  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..>7372  
/gene="NCOR2"  
/locus\_tag="HOM4640"  
ORIGIN  
Query Match : 60.0%; Score 5140.2; DB 10; Length 7372;  
Best Local Similarity 69.9%; Pred. No. 0;  
Matches 5151; Conservative 0; Mismatches 2221; Indels 0; Gaps 0;

OY		AATGTGGGCTCCACA CAGCTTGTGGCA GACGTGTGAAGGCCACTGAACCCCGCTAACCCG	61
Db	1	ATGTGGGATTCACA CAGCTGTGGCA GACGTGTGAAGGCCACTGAACCCCGCTAACCCG	60
OY		CCCCACAGCCTTT CCTACCGA GTAGATCGCCCGGACGCACA CGGAGTGCGGCTCTG	121
Db	61	CCCCACAGCCTTT CTACCGAG TGAATCGCCCGGACGCACA CGGAGTGCGGCTCTG	120
OY		GAGTAACGACACCA CTCGCGGCA CTATAGCTCTCCACTCTGTGGCGGGCTCATCATCCAG	181
Db	121	GAGTACGAGCACCA CTCGCGGCA CTATAGCTCTCCACTCTGTGGCGGGCTCATCATCCAG	180
OY		CCCCAGCGCGGAGG CCGCTCCCTGCTGTGA GTTCCAGCCCGGGAATGAACCGGTCCAG	241
Db	181	CCCCAGCGCGGAGG CCGCTCCCTGCTGTGA GTTCCAGCCCGGGAATGAACCGGTCCAG	240
OY		GAGCTTCAACTGTGGG CCAAGTCCCA CTCTGCTGACCCTGGGGAAGTCAGAGATG	301
Db	241	NNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	300
OY		GAGTTCAATTGAAGA AGACGCCCTCGGCTAGAGCTGTGCTGACCCCTGCTGGACCG	361
Db	301	NNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	360
OY		TCA CCCCTGTGGCCA CGGAGCCGCTCGGGA TCTGAAGA CTTCACCAAGAACCGTAGC	421
Db	361	NNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	420
OY		CTGACGGGCAAGCT TGAAACCGGTGTCTCCCCCAGCCCCCGCACACTGACCCCTGAAGCTG	481
Db	421	CTGAGGGGCAAGCTT GAACCGGTGTCTCCCCCAGCCCCCGCACACTGACCCCTGAAGCTG	480
OY		GAGCTGTGTGCGGCA CGGCTGTCCAAGAGAGAGCTGATCCAGAA CATGCAACCGCGTGGAC	541
Db	481	GAGCTGTGTGCGG CA CGGCTGTCCAAGAGAGAGCTGATCCAGAA CATGCAACCGCGTGGAC	540
OY		CGAGGATCAACCATG TTAGAGCAGCAGATCTTTAA GCTGAAGAAAGACAGCAACGCTG	601
Db	541	CGAGGATCAACCAT GTTAGAGCAGCAGATCTTTAA GCTGAAGAAAGACCAACAGCTG	600
OY		GAGGAGGAGGCTGCA ACGCCGCGGACTGAGAA CCGCGGTGTCACCGCGCCCATGAG	661
Db	601	GAGGAGGAGGCTG CAACGCGCGGACTGAGAA CCGCGGTGTCACCGCGCCCATGAG	660
OY		TGGAAGCACCGCAGC CTTGTGCAATCTTACGACGAGAA CCGGAGMAAGCTGGAAGCT	721
Db	661	TGGAAGCACCGCA GCCTGTGTGCAATCTTACGACGAGAA CCGGNNNNNNNNNNNNNN	720
OY		GCACTCGGATTTCTG AAAGGCTTGCGGGCCCCAG GGTGAGCTGCGCTGTACAACGAGCC	781
Db	721	NNNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	780
OY		TCCGACACCCGGCAG ATCATGAGAACTCAAATAA ACAGGCGCATCGGAAAGAGCTA	841
Db	781	TCCGACACCCGGCAG ATCATGAGAACTCAAATAA ACAGGCGCATCGGAAAGAGCTA	840
OY		ATCTTTGTA ACTTCAAGAGAGGAATCAGCCTCGGAAA CAATGAGACAGAA GTTCTGCAG	901
Db	841	ATCTTTGTA ACTTCAAGAGAGGAATCAGCCTCGGAAA CAATGAGAGCAAA GTTCTGCAG	900
OY		CGCTATGACGAGCTCA TGAAGGCGCTTGGAAAAA AAGGTGAGCGCATCGAAAA CAACCCG	961
Db	901	CGCTATGACGAGCTC ATGAAGGCGCTTGGAAAAA AAGGTGAGCGCATCGAAAA CAACCCG	960
OY		CGCCGCGCGGCCCAG AGAGACGAAGGTGCGCGAGTACTG AAAAAAGCATTTCCCTGAGATC	1021
Db	961	CGCGCGCGCGGCCC AGAGAGAGCAAGGTGCGCGAGTACTG AAAAAAGCATTTCCCTGAGATC	1020
OY		CGCAAGCAGCGGAGCT GACGAGCGCATGCAAGACAGG GTGGCTCAGCGGGGCA GTGGG	1081
Db	1021	CGCAAGCAGCGGAGCT GACGAGCGCATGCAAGACAGG GTGGCTCAGCGGGGCA GTGGG	1080

[illegible]

[illegible][illegible]

Oy 4382 ACCGCGGCTCCACCACTGGCTCCAAAAAGCAGCGTACGGTCCCTCATCGGCGAGCC 4441  
Db 4381 NNNNNNNNGTCCACCACTGGCTCCAAAAAGCAGCGTACGGTCCCTCATCGGCGAGCC 4440  
Oy 4442 GCGCGGAGGTTTCCCACTGGTGAACCCGCTGATGTGATGTGCGGAGCCCGGCACTGGAA 4501  
Db 4441 GCGCGGAGGTTTCCCACTGGTGAACCCGCTGATGTGATGTGCGGAGCCCGGCACTGGAA 4500  
Oy 4502 CGTGCCTCTACGAGAGAGCCTGAAGAGCCGGCAGGGAACCGCAGCAGCTCGGAGGAC 4561  
Db 4501 CGTGCCTCTACGAGAGAGCCTGAAGAGCCGGCAGGGAACCGCAGCAGCTCGGAGGAC 4560  
Oy 4562 TCCATTGGCGGCGGCGCCCGGTCAATTGTGCTGAGCTGGGTAAGCCGCGCAGAGCC 4621  
Db 4561 TCCATTGGCGGCGGCGCCCGGTCAATTGTGCTGAGCTGGGTAAGCCGCGCAGAGCC 4620  
Oy 4622 CTGACCTTATGAGAGCAAGGGGCAACCTTTGGCGGCACTCTCCACAGAGTTGCGCGTG 4681  
Db 4621 CTAACTTATGAGAGCAAGGGGCAACCTTTGGCGGCACTCTCCACAGAGTTGCGCGTG 4680  
Oy 4682 ACCATGCGGAGCCACGCGCGCTGCAAGAGGCAAGCTTTCGTCCAGCAAGGCAATCC 4741  
Db 4681 ACCAAGCGGAGCCCAAGCGCGCTGCAAGAGGCAAGCTTTCGTCCAGCAAGGCAATCC 4740  
Oy 4742 CAGGACCGAAAGCTGACGTGCAAGCTCGTGAGATCGCCAAAGTCCCGCAGCAGCCGTG 4801  
Db 4741 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4800  
Oy 4802 CCCGAGCAGCAGCCCACTCTCGCCCTATGAGCAGCTGCTTCGGGGCTGAGTGAC 4861  
Db 4801 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4860  
Oy 4862 GTGACCTGTATCGAGGCAATCCCGCTGCGCTTGACCCCACTCCATACCCGCGAC 4921  
Db 4861 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4920  
Oy 4922 ATCCCTCTGAGCGAGCGGCTGCTACTACCTGCGCCGACACTTGAGCCCCCAACCCAC 4981  
Db 4921 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4980  
Oy 4982 TACCGGCACTGTATACCAACCTCACTCATCCGGGCTACCCCGACAGCGGCGCTGGAG 5041  
Db 4981 TACCGGCACTGTATACCAACCTCACTCATCCGGGCTACCCCGACAGCGGCGCTGGAG 5040  
Oy 5042 AACCGGAGACCATCATATGATATACATCACTTGACAGAGATGACCAACAAGGAC 5101  
Db 5041 AACCGGAGACCATCATATGATATACATCACTTGACAGAGATGACCAACAAGGAC 5100  
Oy 5102 ACCGCGATGGCCGAGCGAGCTGATATGCTGAGGGGCTCTCGCCCGCGAGTCTCGCTG 5161  
Db 5101 ACCGCGATGGCCGAGCGAGCTGATATGCTGAGGGGCTCTCGCCCGCGAGTCTCGCTG 5160  
Oy 5162 GCACTCACTAGCTGCGGGTCCCGGAGCATCATGACCTGTCCCAAGTGCAACACTG 5221  
Db 5161 GCACTCACTAGCTGCGGGTCCCGGAGCATCATGACCTGTCCCAAGTGCAACACTG 5220  
Oy 5222 CCGTGTGCTGATGCCCGCAGCAGGAGCAACCGGAGCAACCGGAGTGAACCGCTTGCTAC 5281  
Db 5221 CCGTGTGCTGATGCCCGCAGCAGGAGCAACCGGAGCAACCGGAGTGAACCGCTTGCTAC 5280  
Oy 5282 CTCCCGACCGGCGCCAGAGCTTTCAGAGCGGCAAGAGCTGCCACTTCCCAAGAG 5341  
Db 5281 CTCCCGACCGGCGCCAGAGCTTTCAGAGCGGCAAGAGCTGCCACTTCCCAAGAG 5340  
Oy 5342 GGTCCAAACACTTGAACAAAACCAACCAAGTCTCTGTCCGAGCGGAGGGAACCGG 5401  
Db 5341 GGTCCAAACACTTGAACAAAACCAACCAAGTCTCTGTCCGAGCGGAGGGAACCGG 5400  
Oy 5402 GATCGAGAGCGGAGCCGGGATCGGAGCGGAGAAAGTCCATCTCAAGTCCACAGCAG 5461  
Db 5401 GATCGAGAGCGGAGCCGGGATCGGAGCGGAGAAAGTCCATCTCAAGTCCACAGCAG 5460

Oy 5462 GTGAGAGCAGCAACCATCTGAGAGCTTGATACAGAGCAGAGAGCGGCAAGC 5521  
Db 5461 GTGAGAGCAGCAACCATCTGAGAGCTTGATACAGAGCAGAGAGCGGCAAGC 5520  
Oy 5522 AGCGCGGGGGGTGGGGGAGAGAGCGCGCCCGCTCCACTCCATGGCCACAGAGAC 5581  
Db 5521 AGCGCGGGGGGTGGGGGAGAGAGCGCGCCCGCTCCACTCCATGGCCACAGAGAC 5580  
Oy 5582 TCGCCATCTCCCTCGAGACCGAGATGCCCTCAGCAGAGACCAAGTGTCTTACAAAC 5641  
Db 5581 TCGCCATCTCCCTCGAGACCGAGATGCCCTCAGCAGAGACCAAGTGTCTTACAAAC 5640  
Oy 5642 ACAAGCATGAAGGATATATCAACCGCTGTGAGAGCCAGCAAGCCACGCTCTGAGTCC 5701  
Db 5641 ACAAGCATGAAGGATATATCAACCGCTGTGAGAGCCAGCAAGCCACGCTCTGAGTCC 5700  
Oy 5702 AACTCACACTCCTCAACCGGTTGCGCCAGCTGCAATTCCACCTGGACCAACCACTGCCA 5761  
Db 5701 AACTCACACTCCTCAACCGGTTGCGCCAGCTGCAATTCCACCTGGACCAACCACTGCCA 5760  
Oy 5762 CTGGCGGCAACCTTCGATGGGATCTAACCTTACCTCATGAGCCCGTCTTGCGCCAAAG 5821  
Db 5761 CTGGCGGCAACCTTCGATGGGATCTAACCTTACCTCATGAGCCCGTCTTGCGCCAAAG 5820  
Oy 5822 GAGGCCCCCGGGGTGCGCCGCGCAGAGCGGCGCCGAGCAGACCTGGCATTGCTTCTC 5881  
Db 5821 GAGGCCCCCGGGGTGCGCCGCGCGCAGAGCGGCGCCGAGCAGACCTGGCATTGCTTCTC 5880  
Oy 5882 GCAAGAGCCCGAGCGGCTCGGGGCTGGAGCGCGCTCTCCCCAGCAAGGGCTCGGAG 5941  
Db 5881 GCAAGAGCCCGAGCGGCTCGGGGCTGGAGCGCGCTCTCCCCAGCAAGGGCTCGGAG 5940  
Oy 5942 CCGCGGCCCTTATGCTCTCTGTCTTGCGCAGCAGCAATCGCCGCAACCCCTGCGAAG 6001  
Db 5941 CCGCGGCCCTTATGCTCTCTGTCTTGCGCAGCAGCAATCGCCGCAACCCCTGCGAAG 6000  
Oy 6002 AACTCGCACTTCAACACGCGAGCGCGGAGCCCGCGCGGCACTGGCTCGGCTCGGAC 6061  
Db 6001 AACTCGCACTTCAACACGCGAGCGCGGAGCCCGCGCGGCACTGGCTCGGCTCGGAC 6060  
Oy 6062 CCGGACCGGGAAGAGCTCAAGTAAACCTTTTCATCCAGGAACTGGAACTCGGTTCT 6121  
Db 6061 CCGGACCGGGAAGAGCTCAAGTAAACCTTTTCATCCAGGAACTGGAACTCGGTTCT 6120  
Oy 6122 CTGGTTACACAGGAGGACTACAGCCCGAAGGGGTGAGCCCGTCAAGCCTGTGAGC 6181  
Db 6121 CTGGTTACACAGGAGGACTACAGCCCGAAGGGGTGAGCCCGTCAAGCCTGTGAGC 6180  
Oy 6182 TCAACCACTTGAGCCACGACAGAGGGGCTCCCAAGCACTGGAAAGACTGACAGAGC 6241  
Db 6181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6240  
Oy 6242 CACTGAGAGGGGAGCTGGGCGCCAGAGCGAGCGCCCGTGAAGCTTGGGGGAGGCG 6301  
Db 6241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6300  
Oy 6302 GCCCACTCCCAACTTGGCGCGCTGCTGAGAGCAGCCCTGTCTCAAGCCGCTGTCTC 6361  
Db 6301 GCCCACTCCCAACTTGGCGCGCTGCTGAGAGCAGCCCTGTCTCAAGCCGCTGTCTC 6360  
Oy 6362 CAGAGCGGCCCAAGGGGTCAAAAGTCAACAGCGGGTGTCACTTGGCCAGACATTCAGT 6421  
Db 6361 CAGAGCGGCCCAAGGGGTCAAAAGTCAACAGCGGGTGTCACTTGGCCAGACATTCAGT 6420  
Oy 6422 GAGGTCAATCAAGAGATACACCGGCAACACCAAGCAAGCTCAAGGCAACCCCTGCGC 6481  
Db 6421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6480  
Oy 6482 GCCCCCTCTACTCTCTTCTTGGGGCGAGCTGCCCGTCTGGAACCTCGCGGCCACCC 6541  
Db 6481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6540  
Oy 6542 AGTGAACCTTACTCTCCGCCCCGAGCAATGGTCCCGGCGCTGCTCCCCCAAGC 6601



[illegible]

TITLE	Adams,M.D. and Cargill,M.
JOURNAL	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
PUBMED	Science 302 (5652), 1960-1963 (2003)
REFERENCE	14671302
AUTHORS	2 (bases 1 to 6836)
TITLE	Clark,A.G., Gnanowsk,i,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tannenbaum,D.M., Civejello,D.R., Lu,F., Murphy,B., Reriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
JOURNAL	Direct Submission
COMMENT	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
FEATURES	This sequence was made by sequencing genomic exons and ordering them based on alignment.
source	Location/Qualifiers
gene	1..6836
	/organism="Pan troglodytes"
	/mol_type="Genomic DNA"
	/db_xref="taxon:9598"
	<1..>6836
	/gene="NCR2"
	/locus_tag="HCM4640"
ORIGIN	
Query Match	49.2%; Score 4216.6; DB 10; Length 6836;
Best Local Similarity	62.1%; Pred. No. 0;
Matches 4243;	Conservative 0; Mismatches 2593; Indels 0; Gaps 0
QY	79 CCCAGTGCAGATCGCCCGGACGACACGAGCGTGGGCTCTGAGTACCAAGCACCATC 138
Db	1 CCAGTGCAGATTGCCGACGACACGAGCGTGGGCTCTGAGTACCAAGCACCATC 60
QY	139 CCGGCACATATGCTCTCCACCTGTGCGCGGGGTCCATATACACCCACGGGGGAGGCC 198
Db	61 CCGGCACATATGCTCTCCACCTGTGCGCGGGTCCATATCACCCACGGGGGAGGCC 120
QY	199 CTCCTGCTGTCGTAGGTTCCAGCCCGGGAAATGAAACGATCCAGAGCTCAACTGCGGCC 258
Db	121 CTCCTGCTGTCGTAGTTCCAGCCCGGGAAATGAACGATCCAGAGCTCAACTGCGGCC 180
QY	259 AGAGTCCCACTATACCTGCGGAGCTGGGGAACTCAGATGAGTTCAATTGAAGCAA 318
Db	181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
QY	319 GGGCCCTCGGCTAGAGCTGCTGCTGACCCCTGTGCTGACCGTCAACCTGCTGGCCAC 378
Db	241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
QY	379 GGGCCAGCTCGGGGATCTGAAGACTCACCAAGAACCTGAGCTGACGGGCAAGCTGGA 438
Db	301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360
QY	439 ACCGGTGTCTCCCCCAGACCCCGGACACTGACCTGAGCTGAGCTGTGTCGCGCACG 498
Db	361 ACCGGTGTCTCCCCCAGACCCCGGACACCGAACCTGAGCTGAGTGGTGTGCGCGCG 420
QY	499 GCTGTCCAAAGAGAGCTGATCCGAAACATGACCCGGTGTGACCCGAGAGATCAACATGCT 558
Db	421 GCTGTCCAAAGAGAGCTGATCCGAAACATGACCCGGTGTGACCCGAGAGATCAACATGCT 480
QY	559 AAGAGCAGAGATCTCTAAGCTGAAGAAAGACAGCAACAGCTGAGAGAGAGGCTGCCAA 618
Db	481 AAGAGCAGAGATCTCTAAGCTGAAGAAAGACAGNNNNNNNNNTTGAAGAGAGGCTGCCAA 540
QY	619 GCGGCCCAAGCTGAGAAACCCGTGTCAACGCGGCCCATGAGTGAAGACCGCAGCTT 678
Db	541 GCGGCCCAAGCTGAGAAAGCCCGTGTCAACGCGGCCCATGAGTGAAGACCGCAGCTT 600
QY	679 GGTGCAATCATTTACGACGAGAACCGGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 738
Db	601 GGTGCAATCATTTACGACGAGAACCGGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 660

OY	739	AGCCTGGGCGCCCAAGTGTGAAGTCGCCCTGTACAACAGCCCTTCCACACCOCGCAGTA	798
Db	661	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCCCTGTACAACAGCCCTTCCACACCOCGCAGTA	720
OY	799	TCATGAGAATCAATAAATAAACACAGGCGCATGCGGAAAGAAAGCTAATTCTTGACTTTCAAG	858
Db	721	TCAATGAGAATCAATAAATAAACACAGGCGCATGCGGAAAGAAAGCTAATTCTTGACTTTCAAG	780
OY	859	GAGGAATCACGCTCGGAAAACAATGGAACAAGATTCTGCCAGCGCTATGACCAAGCTCAT	918
Db	781	GAGGAATCACGCTCGGAAAACAATGGAACAAGATTCTGCCAGCGCTATGACCAAGCTCAT	840
OY	919	GGAGGCCCCCTGGAABAAAAAGGTGAGCGGCATGAAAACAACCCCGCGCGGCGCTCAAGA	978
Db	841	GGAGGCCCCCTGGAABAAAAAGGTGAGCGGCATGAAAACAACCCCGCGCGGCGCTCAAGA	900
OY	979	GAGCAAAGTGGCGCGAGTACTAGAAAACAGTTCCTCTTAGATTCGCGAAGCGCGAGCT	1038
Db	901	GAGCAAAGTGGCGCGAGTACTAGAAAACAGTTCCTCTTAGATTCGCGAAGCGCGAGCT	960
OY	1039	GCAGAGCGCATGCAAGACAGAGGTGGGCGACGCGGGCAGTGGGCTGTTCATGTGGCGCG	1098
Db	961	GCAGAGCGCATGCAAGACAGAGGTGGGCGACGCGGGCAGTGGGCTGTTCATGTGGCGCG	1020
OY	1099	CCGCAAGCGACACAAGGCTGTCAAGATCATGTGGCTCTTCAGACAGGAACAATTGGA	1158
Db	1021	CCGCAAGCGACACAAGGCTGTCAAGATCATGTGGCTCTTCAGACAGGAACAATTGGA	1080
OY	1159	GAGCAGATGGCGCGAGCTGGCGCGTAGTCCGCGCCATGCTGACAGCGCTGACAGCGCG	1218
Db	1081	GAGCAGATGGCGCGAGCTGGCGCGTAGTCCGCGCTTGTGCTGACAGCGCGCGACAGCGCG	1140
OY	1219	CATCAAGTTCATCAACATGAACGGGCTTTATGGCCGACCCCATGAAAGTGTACAAGAACG	1278
Db	1141	CATCAAGTTCATCAACATGAACGGGCTTTATGGCTGACCCCATGAAAGTGTACAAGAACG	1200
OY	1279	CCAGGCTCATGAACATGTGTGAGTGAAGCAGAGAGAAAGGACCTTTCGGGAGAAATTCATGCA	1338
Db	1201	CCAGGCTCATGAACATGTGTGAGTGAAGCAGAGAGAAAGGACCTTTCGGGAGAAANNNNNNN	1260
OY	1339	GCATCCCAAGAACTTTGGCTGTGATGCATCATTCCTGGAAGAGAAACAAGTGGCTGAGTG	1398
Db	1261	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1320
OY	1399	CGTCCCTCATTAATTACTACCTGAAGAAATGAGAACTATAAGACCTGTGTAGACGAG	1458
Db	1321	CGTCCCTCATTAATTACTACCTGAAGAAATGAGAACTATAAGACCTGTGTAGACGAG	1380
OY	1459	CTATTCGCGCGCGCGCAAGAGCAGCAGCAACAACAGACAGCAGCAGCAGCAGCAGCA	1518
Db	1381	CTATTCGCGCGCGCGCAAGAGCAGCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1440
OY	1519	GCAGCAGACAGCGCCAATGCCCGCAGACGCCAGAGAGAAAGATGAGAAAGAGAAAGGA	1578
Db	1441	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1500
OY	1579	AAGAGAGCGAGAGAAAGAGAGAGAACCGGAGGTGAGAAACGACAAGAAACCTCCT	1638
Db	1501	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1560
OY	1639	CAAGAGAGAACAAGACGACACTCAAGGAGAGACAACAGACGAGAAGAGAGCTGTGACCTC	1698
Db	1561	NNNGGAGAGAACAAGACGACACTCAAGGAGAGACAACAGACGAGAAGAGAGCTGTGACCTC	1620
OY	1699	CAAGAGCGGCAAACTGTCCAAACAGCAGGAGAAACGCAAAAGCGGCATCAACCGCTCAT	1758
Db	1621	CAAGAGCGGCAAACTGTCCAAACAGCAGGAGAAACGCAAAAGCGGCATCAACCGCTCAT	1680
OY	1759	GGCTTAATGAGCGCAACAGCAGAGAGGCAATCACCCCGCAGAGAGCGCGGAGCTGGCCTC	1818
Db	1681	GGCTTAATGAGCGCAACAGCAGAGAGGCAATCACCCCGCAGAGAGCGCGGAGCTGGCCTC	1740
OY	1819	CATGAGCTGATGAGATTCTCGCTGACAAGAAAGAAATGGAALCAGCCAGAAAG	1878

Db	1741	CATGAGCTGATGAAGTTCTCGCTGAAAGAAAGAAATGAAACACGCAAGAAAG	1800
Qy	1879	TCCTCTGGAAACAAGCGCGCAACTGGTGGGCATCGCCGATGGTGGGCTCCAAGCTGT	1938
Db	1801	TCTCTGGAACAAGCGCGCAACTGGTGGGCATCGCCGATGGTGGGCTCCAAGCTGT	1860
Qy	1939	GTGCACTGTGAAGACTTCTTACTTCAACTAGAGAGGCAAGACTCGATGATCTT	1998
Db	1861	GTCCACATGTGAAGACTTCTTACTTCAACTAGAGAGGCAAGACTCGATGATCTT	1920
Qy	1999	GCACACAGCAAGCTGAAGATGGAAGAGAGAGGAAAGCGCGGAGAAAGAAAGAAAGC	2058
Db	1921	GCACACAGCAAGCTGAAGATGGAAGAGAGAGGAAAGCAGAGAGAAAGAAAGAAAGC	1980
Qy	2059	GCCGCGCGCGGCGCAGCAGAGAGGCTGCATTCGCCCGTGGTGAAGATAGAGATGA	2118
Db	1981	GCCGCGCGCGGCGCAGCAGAGGCTGCATTCGCCCGCTGGTGAAGATAGAGATGA	2040
Qy	2119	GGCGTCGGGCGTGAGCGGAAATGAGAGAGATGGTGAAGAGCTTGAAAGCTTACATGC	2178
Db	2041	GGCGTCGGGCGTGAGCGGAAATGAGAGAGATGGTGAAGAGCTTGAAAGCTTACATGC	2100
Qy	2179	CTCTGGGAATGAGGTGCCAGAGGGGAAATGCAGTGGCCAGCCACTGTCAACACAGCTC	2238
Db	2101	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	2160
Qy	2239	AGACACGAGAGACTCCCTCTCCTCACTGAGGCGCGCAAGACAAGGCAAGATG	2298
Db	2161	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	2220
Qy	2299	GCCCAAGCCCCAGACCACTTGAGGCGCGCAGCGGCCACCCAGGCCACCACCCCACC	2358
Db	2221	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	2280
Qy	2359	ACGGAAGACATCCGGGGCCCCATTAGGCCACCCGGGCTCTGAAGCACCGAGCCCC	2418
Db	2281	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	2340
Qy	2419	TAGCCCCCACAAGCACCCCATCGCCTCTGCACCTCTCTGTGTGCCAAGAGAGA	2478
Db	2341	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	2400
Qy	2479	GAAGAGAGAGAGACCGCAGACAGGCCCCAGTGAAGAGGGGAGAGAGACGCCCC	2538
Db	2401	NNAGAGAGAGAGACCGCAGACAGGCCCCAGTGAAGAGGGGAGAGAGAGAGCCCCC	2460
Qy	2539	CGCGGCTGAAGAGCTGGCACTGACACAGGAAAGCCGAGAGGCCGTCAAGAGCGAGTG	2598
Db	2461	CGCANNAGAGAGCTGGCACTGACACAGGAAAGGCCAGAGAGGCCGTCAAGAGCGAGTG	2520
Qy	2599	CAGCGAAGAAAGCGAAGAGGGGGCGCGCAAGGGCAAGGAAGCGGAGGGCGCTAGAGGCAC	2658
Db	2521	CACGGAAGAAAGCGAAGAGGGGGCGCGCAAGGGCAAGGAAGCGGAGGGCGCTAGAGGCTAC	2580
Qy	2659	GGCGGAGGGGCGCTCAAGGCAAGAAAGAAAGAGGGCGGAGGGCGCAGGGCCACACTGC	2718
Db	2581	GGCGGAGGGGCGCTCAAGGCAAGAAAGAAAGAGGGGGGAGGGCGCAGGGCCACACAGC	2640
Qy	2719	CAAGAGCTCGGGCGCCCCCAAGACAGCGACTTCAGTGTCTGACGTGCAAGTGCAGAGT	2778
Db	2641	CAAGAGCTCGGGCGCCCCCAAGACAGCGACTTCAGTGTCTGACGTGCAAGTGCAGAGT	2700
Qy	2779	GGAAGAGCCGAGGGCGCGCGCAAGAAACGGGCTGCTGTCCCAAGGCCAGGCTCTCTCAC	2838
Db	2701	GGAAGAGCCGAGGGCGCGCGCAAGAAACGGGCTGCTGTCCCAAGGCCAGGCTCTCTCAC	2760
Qy	2839	CCCGACTGGGCAACCCCGGGGCAATGGCTCAACCCAGAAACCACTGGACTTGAAGCACT	2898
Db	2761	CCCGACTGGGCAACCCCGGGGCAATGGCTCAACCCAGAAACCACTGGACTTGAAGCACT	2820
Qy	2899	GAAGCAGCAGGGCTGCATCCCCCCCATCGAGTCAACCAAGTCCATGAGCCCCCG	2958

Db 2821 GAAGACGAGCGGCTGCCATCCCCCATCCAGGTCAACCAAGTCCATGAGCCCCCG 2880  
QY 2959 GGAGAGCGAGCTCCACCAAGCCAGTCCCCAGGCCCAACGCGCAACGCAAAACCTGCA 3018  
Db 2881 GGAGAGCGAGCTCCACCAAGCCAGTCCCCAGGCCCAACGCGCAACGCAAAACCTGCA 2940  
QY 3019 GCCGAGAGCGAGCGCCCTCAGAGCCTGAGCAGAGCCCCCGGGGCAAGAGAGAGCC 3078  
Db 2941 GCCGAGAGCGAGCGCCCTCAGAGCCTGAGCAGAGCCCCCGGGGCAAGAGAGAGCC 3000  
QY 3079 GGCAACCCCGCCGCAAGAGAGCCTTCGAGCCGAGGCCCAGAGAGCTGCTGGGACCC 3138  
Db 3001 GGCAACCCCGCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
QY 3139 CCTTGTCTGAGCTTCGGGCTGCTTCCCGTGGCCCGCCGAGGTGATCAAGGCTTC 3198  
Db 3061 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3120  
QY 3199 CCGCAGTGGCCCGGAGCCCTCAGCTTCTCTAGCTCCAGCTGTCAACCACTGCCCC 3258  
Db 3121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3180  
QY 3259 GGGCTTCATGACATGCGCCGCGCTCTGCGCGGCCCAACCAATCTTCAACCCGCC 3318  
Db 3181 GGGCTTCATGACATGCGCCGCGCTCTGCGCGGCCCAACCAATCTTCAACCCGCC 3240  
QY 3319 TCCCTCATCTCTCTGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3378  
Db 3241 TCCCTCATCTCTCTGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300  
QY 3379 CCAAGAGATGTCGATCAGCTCAGCTCCGCTACTAGAGAGATGCCAAGGCCCGGTGAG 3438  
Db 3301 CCAAGAGATGTCGATCAGCTCAGCTCCGCTACTAGAGAGATGCCAAGGCCCGGTGAG 3360  
QY 3439 CCCTGTACCATGAGGAGCTGCGCTGCGCATGAGACCCCAAAAGCTGACCCCTTCA 3498  
Db 3361 CCGCTGACCATGAGGAGCTGCGCTGCGCATGAGACCCCAAAAGCTGACCCCTTNN 3420  
QY 3499 AATTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3558  
Db 3421 AATTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3480  
QY 3559 GCCCAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3618  
Db 3481 GCCCAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540  
QY 3619 AAGCATCACCAAGGAGATTCCAGACACAGGAGTCCCTGGAACAGGCGCATCAATAC 3678  
Db 3541 AAGCATCACCAAGGAGATTCCAGACACAGGAGTCCCTGGAACAGGCGCATCAATAC 3600  
QY 3679 CGGCTTCATCAACCAAGGAGATTCCAGACACAGGAGTCCCTGGAACAGGCGCATCA 3738  
Db 3601 CGGCTTCATCAACCAAGGAGATTCCAGACACAGGAGTCCCTGGAACAGGCGCATCA 3660  
QY 3739 CATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3798  
Db 3661 CATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720  
QY 3799 CCAAGTCACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3858  
Db 3721 CCAAGTCACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780  
QY 3859 GACCCAGTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3918  
Db 3781 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3840  
QY 3919 CCGCAAGCGCACTATGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3978  
Db 3841 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3900  
QY 3979 CATCGAAGGTCTCATGAGGCTGTCATCCCGCGAGAGAGAGAGAGAGAGAGAGAGAG 4038  
Db 3901 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3960

QY 4039 AGAGACGACACATCCGGGGTTCATCAACAAGAGATCCCTGGTCTAGTGAGAGCC 4098  
Db 3961 AGAGACGACACATCCGGGGTTCATCAACAAGAGAGATCCCTGGTCTAGTGAGAGCC 4020  
QY 4099 ACAGAGAGACTAGCTGCTGGAGAGCGCAAGCTCTTAAAGCGAGAGAGAGAGAGAG 4158  
Db 4021 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4080  
QY 4159 CCAACCGCTCAGCGGAGCTTACAGAGGCTTACAGAGAGAGAGAGAGAGAGAGAGAG 4218  
Db 4081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4140  
QY 4219 GCTGAAGCGGCGCATGAGAGGCTGTGAGCAAGGTGAAGAGAGAGAGAGAGAGAG 4278  
Db 4141 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4200  
QY 4279 TGAAGTCCGCGGAGAGAGCTGCGGCAAGCCGAGCTGCGCCCGCGCGCT 4338  
Db 4201 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4260  
QY 4339 CAAAGAGGCTCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4398  
Db 4261 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4320  
QY 4399 TGGCTCAAAAGACAGAGAGTACGCTCCCTCATCGGAGAGAGAGAGAGAGAGAG 4458  
Db 4321 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4380  
QY 4459 CGTGACCCGCTGATGATGAGCGAGCGCCGAGCACTGGAACGTGCTGCTACAGAGA 4518  
Db 4381 CGTGACCCGCTGATGATGAGCGAGCGCCGAGCACTGGAACGTGCTGCTACAGAGA 4440  
QY 4519 GAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4578  
Db 4441 GAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4500  
QY 4579 CCGGCTCATTTGCTGAGCTGAGGTAAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 4638  
Db 4501 CCGGCTCATTTGCTGAGCTGAGGTAAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560  
QY 4639 CGGAGGACCCCTTTCGAGGCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4698  
Db 4561 CGGAGGACCCCTTTCGAGGCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4620  
QY 4699 GCGGAGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4758  
Db 4621 GCGGAGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4680  
QY 4759 GTGAGAGCTGTGAGATGCGCAAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 4818  
Db 4681 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4740  
QY 4819 CCGCATCTCGCCCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4878  
Db 4741 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4800  
QY 4879 CCAATCCCTGAGCTTGAACCCAGCTTCAATCCCGGAGATCCCTTGAAGAGAGAG 4938  
Db 4801 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4860  
QY 4939 CGCTGCTTACATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4998  
Db 4861 CGCTGCTTACATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4920  
QY 4999 ACCCTACCTCATCCGAGGCTTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5058  
Db 4921 ACCCTACCTCATCCGAGGCTTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4980  
QY 5059 CAATGACTCATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5118  
Db 4981 CAATGACTCATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5040

QY 5119 AGCTGATATGCTGAGGGGCTCTCGCCCCGAGTCTCGTTGGCACTCACTACGCTGC 5178  
| | | | |  
Db 5041 AGCTGACATGTGAGGGGGCTCTCGCCCCGAGTCTCGTTGGCACTCACTACGCTGC 5100  
| | | | |  
QY 5179 GGGTCCCGGAGGCAATCGACTGTGCCAATGGCAACCTGCTGTGCTGTGCCCCC 5238  
| | | | |  
Db 5101 GGGTCCCGGAGGCAATCGACTGTGCCAATGGCAACCTGCTGTGCTGTGCCCCC 5160  
| | | | |  
QY 5239 GACACCAAGGCAACCCAGGCCACCGCATGAGCCGCTTGCTCACTCCACCGCCGCCA 5298  
| | | | |  
Db 5161 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5220  
| | | | |  
QY 5299 GCCCTTGAGAGCCCGCACAGAGCTCCCACTCTGCCAGAGGTCCAACTGATC 5358  
| | | | |  
Db 5221 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5280  
| | | | |  
QY 5359 AAAACCAACCAACGATCTGCTCGAGGGGAGGAGCGGGATTCAGAGACGGGAGCCG 5418  
| | | | |  
Db 5281 NNAACCAACCAACGATCTGCTCGAGGGGAGGAGCGGGATTCAGAGACGNNNNNN 5340  
| | | | |  
QY 5419 GGATCGGAGGAGCGGAAAAGTCCATCTCTCACTCAACAGCGGTGAGACGACCCAT 5478  
| | | | |  
Db 5341 NNNNNNNNAGCGGAAAAGNCATCTCTCACTCAACNNNNNNGTGAGACGCCCCAT 5400  
| | | | |  
QY 5479 CTGAGACCTGTTACAAGACAGAGAGCGGACAGCGGACGCGCGGGGTGGGG 5538  
| | | | |  
Db 5401 CTGAGNNNNNNGTACAAGACAGAGAGCGGACAGCGGTGGGGGTGGGG 5460  
| | | | |  
QY 5539 CAGCAGAGCGCGCGCTCCCACTCCCATGCTCCCAAGCACTGCGCATCTCCCTCG 5598  
| | | | |  
Db 5461 CAGCAGAGCGCGCGCTCCCACTCCCANNNCCACAGCACTGCGCATCTCCCTCN 5520  
| | | | |  
QY 5599 GACCCAGATGCTCTCAGCAGAGACCCAGTGTCTTCAACAAGCATGAGGGTAT 5658  
| | | | |  
Db 5521 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5580  
| | | | |  
QY 5659 CATCACCGCTGTGAGAGCCAGAGCCACGGTCTGAGGTCAACTCTCTCACTC 5718  
| | | | |  
Db 5581 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5640  
| | | | |  
QY 5719 GCTTGCCAGCTGCACTTCCCACTGCAACCACTGCGCATGCGGAGGACCTCGA 5778  
| | | | |  
Db 5641 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5700  
| | | | |  
QY 5779 TGGGCTTACCTCACTCATGAGCCCGTCTTGCTGCCAAGAGAGCCCGCGGTGC 5838  
| | | | |  
Db 5701 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5760  
| | | | |  
QY 5839 CCGGCCAGAGCGGCCCCGAGACAGACGCGCATGCTTCTCGCAAGCCCCAGCCG 5898  
| | | | |  
Db 5761 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5820  
| | | | |  
QY 5899 CTCCGGGCTGAGCGCGCTCTCCCGCAGAGGGCTCGAGCCCGCGCTGTAGTCC 5958  
| | | | |  
Db 5821 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5880  
| | | | |  
QY 5959 TCTGCTCTGCGCAGCCACCATGCGCCGACCCCTGCGAAGAACTTGACACTCACA 6018  
| | | | |  
Db 5881 CCCGCTCTGCGCAGCCACCATGCGCCGACCCCTGCGAAGAACTTGACACTCACA 5940  
| | | | |  
QY 6019 CGCCAGCCCGGAGCCCGCGCGCCACTGCTCGGCTTGCGACCGCGAAGAAAGC 6078  
| | | | |  
Db 5941 CGCCAGCCCGGAGCCCGCGCGCCACTGCTCGGCTTGCGACCGCGAAGAAAGC 6000  
| | | | |  
QY 6079 TCAAGTAAACCTTTCCATCAGAGAACTGGAATCGGTTCTTGCGGTACACAGCGAG 6138  
| | | | |  
Db 6001 TCAAGTAAACCTTTCCATCAGAGAACTGGAATCGGTTCTTGCGNNNNNNNNNNNN 6060  
| | | | |  
QY 6139 CAGCTACAGCCCGAAGGGGTGAAGCCGCTGAGCTCAACAGTCTGACCCA 6198  
| | | | |  
Db 6061 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6120  
| | | | |  
QY 6199 CGACAGAGGGCTCCCAAGCACCTGGAAGAGCTCGACAAGGCCACTTGAGAGGGGAGCT 6258  
| | | | |

Db 6121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6180  
| | | | |  
QY 6259 GCGGCCAAGAGCAGGCCCCGTGAAGCTTGGCGGGAGGCCGCACTGCCACACT 6318  
| | | | |  
Db 6181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6240  
| | | | |  
QY 6319 GCGGCGCTGCTGAGAGCAGCCCTCGTCAAGCCGCTGCTCAAGCCGCCAGGGGT 6378  
| | | | |  
Db 6241 GCGGCGCTGCTGAGAGCAGCCCTCGTCAAGCCGCTGCTGCTGAGACCGGCCAGGGGT 6300  
| | | | |  
QY 6379 CAAAGTCAACAGCGGGGTGTCACTTGGCCGACATCAGTGAAGTATCAACAGGA 6438  
| | | | |  
Db 6301 CAAAGTCAACAGCGGGGTGTCACTTGGCCGNNNNNNNNNNNNNNNNNNNNNNNNNN 6360  
| | | | |  
QY 6439 CTACACCCGGACACCAACAGAGCTAGAGGCAACCTGCCCCCTCTACTCCTT 6498  
| | | | |  
Db 6361 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6420  
| | | | |  
QY 6499 CCTGGGGCCAGCTGCCCTGCTGAGACTCCGCGCCCAACCACTGACTCTGCC 6558  
| | | | |  
Db 6421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6480  
| | | | |  
QY 6559 GCCCCCGGACCATGTGCTCCCGGCGCTGCTCCCCCAAGCGAAGGGGCAAGGTG 6618  
| | | | |  
Db 6481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6540  
| | | | |  
QY 6619 TCCAGAGCCAAACAAGAGCTGCTTGGGTGTGTTGAGAGCGGTATTAACCTGTGTC 6678  
| | | | |  
Db 6541 TCCAGAGCCAAACAAGAGCTGCTTGGGTGTGTTGAGAGCGGTATTAACCTGTGTC 6600  
| | | | |  
QY 6679 CCAACCGGAGGCAATGACGAGGCAAGGCACTCCGAGTCTGTGTAACCGCTGCTGA 6738  
| | | | |  
Db 6601 CCAACCGGAGGCAATGACGAGGCAAGGCACTCCGAGTCTGTGTAACCGCTGCTGA 6660  
| | | | |  
QY 6739 CCGGATGAGGGAACAGAGAGGCCAGAGATGAGTCTCAAGTCTCAAGGCAACAG 6798  
| | | | |  
Db 6661 CCGGATGAGGGAACAGAGAGGCCAGAGATGAGTCTCAAGTCTCAAGGCAACAG 6720  
| | | | |  
QY 6799 CGAGCGGCGAGCTTCTTCAAGAGCTGACGAGAGCAATCCGCGATGCTCAAGTCCA 6858  
| | | | |  
Db 6721 CGAGCGGCGAGCTTCTTCAAGAGCTGACGAGAGCAATCCGCGATGCTCAAGTCCA 6780  
| | | | |  
QY 6859 GAAGCAAGATCAACAAGAGCTGAACCCACAACCGAATGAGCTGATATCA 6914  
| | | | |  
Db 6781 GAAGCAAGATCAACAAGAGCTGAACCCACAACCGAATGAGCTGATATCA 6836  
| | | | |  
  
RESULT 3  
AY412688  
LOCUS 7013 bp DNA linear GSS 16-DEC-2003  
DEFINITION Mus musculus NCOR2 gene, VIRUAL TRANSCRIPT, partial sequence,  
ACCESSION AY412688  
VERSION AY412688.1 GI:39768653  
KEYWORDS  
SOURCE  
ORGANISM Mus musculus (house mouse)  
  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Muridae; Mus.  
  
REFERENCE  
1 (bases 1 to 7013)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tenenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,  
Adams,M.D., and Cargill,M.  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
2 (bases 1 to 7013)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tenenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,

Ferriere, S., Wang, G., Zheng, X.H., White, T.J., Sinsky, J.J.,  
Adams, M.D. and Cargill, M.  
Direct Submission  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
COMMENT Location/Qualifiers  
FEATURES  
source 1..7013  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
<1..>7013  
/gene="NCOR2"  
/locus\_tag="HMC640"

## ORIGIN

Query Match 38.7%; Score 3315.2; DB 10; Length 7013;

Best Local Similarity 58.3%; Pred. No. 0;

Matches 4299; Conservative 0; Mismatches 2711; Indels 365; Gaps 14;

```
QY 2 ATGTGGGGCTCCACACAGCTTTGTGGACAGAGTGGAGGGCCACTGAGCCCGCTACCG 61
DB 1 ATGTGAGATCCACACAGCTGTGGACAGACATGGGGGCTGCTGAGCCCGCTACCGA 60
QY 62 CCCACAGCCTTCTACCCAGTGCAGATGGCCGGAGCAGACGGAGCGTGGGCTCTG 121
DB 61 CCCATGGATCTCTTACCCGGTGCAGATAGCCGGTCCACAGGAGTGGGGCTGCTT 120
QY 122 GAGTACACAGACCACTCCGCGACTATGCTCTCCACTGTGCGCGGGCTTCATCTCA 181
DB 121 GAGTACCAACACACACCCCGTGACTACACTCACTGTCACTCCGGTTCATCTCA 180
QY 182 CCCACGGCGGAGGGCTCTCTGTGTCTGAGTTCAAGCCGGGAATGAACGCTCCAG 241
DB 181 CCACAGAGAGGGGCGCTCTCACTGTGTGAGTTCCAGCTGGAGTGAAC-GNNNNNN 239
QY 242 GAGCTCACCTGGGCGAGAGTCCACTCATCTCCCGAGCTGGGGAAGTACAGATG 301
DB 240 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 299
QY 302 GAGTTCAATTGAAGCAAGCGCCCTCGGCTAGAGCTGCTGACCCCTGCTGCGACG 361
DB 300 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 359
QY 362 TCAACCTCTGCTGGCCACGGGCCAGCTGGGGATCTGAAGACTCAACAGACCGTAG 421
DB 360 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 419
QY 422 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCGAGCCCCCGGACACTGACCTGAGCT 481
DB 420 CTGGCAGGCAAGCTGGAAGCTGTGTCACTCCAGTCCCGGAGCTGACCTGTAGCTTA 479
QY 482 GAGCTGTGCTCCGCAAGCTGTCTCAAGAGAGAGCTATCCAGAACTGACCCGCTGAC 541
DB 480 GAGCTGTGCGCACTGTGACTGTCTCAAGAGAGAGCTATCCAGAACTGACCCGCTGAC 539
QY 542 CGAGAGATCACTGTTAGAGCAGAGCTCTTAAGCTGAAGAAAGACAGCAACGCTG 601
DB 540 CCGTGAATCACTGTTAGAGCAGAGCTCTCAAGCTGAAGAAAGACAGCAACGCTG 599
QY 602 GAGGAGGAGGCTGCGCAAGCGCCCGGAGCTGAGAAACCGGTGTCAACCGCCGCTGAG 661
DB 600 GAGGAGGAGGCGCCAGAGCGCCGGAACCGGAGAACTGTGTGCGCAACCCATAGAA 659
QY 662 TCGAAGCAACCGAGCTGTGTGAGATCATCTAGACAGAGAACCGAAGAGGCTGAAGCT 721
DB 660 TCGAAGCAACCGAGCTGTGTGAGATCATCTAGAGAGAACCGGNNNNNNNNNNNNNN 719
QY 722 GCACATCGAATTTGGAAGGCTGGGGCCCAAGTGAAGTGTGCGCTGTACAACGAGCC 781
DB 720 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 779
```

```
QY 782 TCCGACACCGGCGATATCATGAAATCAATAAACCAGGCGATGCGGAAGAGCTA 841
DB 780 TCTGACACACGCGCAGTACATGAAACATCAAAATTAACAGGCGATGCGGAAGAGCTG 839
QY 842 ATCTTGAATCTTCAAGAGAGAAATCACGCTCGGAACATGAAAGCAAGTTCTGCGAG 901
DB 840 ATCTTGAATCTTAAAGCGAGAACACAGCGCCGAGAGTGGGAACAGCCCTCTGCGAG 899
QY 902 CGCTATGACAGCTCATGAGAGGCTTGGAAAAAAGGTGAGGGCATCGAAAACAACCCG 961
DB 900 CGCTATGACAGCTCATGAGAGGCTTGGAAAAAAGGTGAGGGCATAGAAACATCCG 959
QY 962 CGCCGCGGCGCAAGAGAGCAAGTGTGCGAGTACTACGAAAGAGATTCCCTGAGATC 1021
DB 960 CGAAGAGGGCCCAAGAGAGCAAGTGAAGGATCTACTAGAGAAACAGTTCCCGAGATC 1019
QY 1022 CGCAGACAGCGAGCTGCGAGAGGCGCATGACAGACAGGCTGGGCCAAGCGGCGAGTGG 1081
DB 1020 CGCAGACAGCGGAGCTGCGAGAGGCGCATGACAGACAGGCTGGGCCAAGCGTGGCAGTGGG 1079
QY 1082 CTGTCAATGTGCGCGCGCCGCGAGCGAGCAGAGTGTCAAGATCATGCAATGCGCTCTCA 1141
DB 1080 CTCTCATGTGCGCTGCGCGAGTGCATGAGTTCGAGATCATGATGCTGTCT 1139
QY 1142 GAGCAGAGAACTGTGAGAAAGCAGATGCGCCAGCTGGCGCTGATCCGCGCATGCTGAC 1201
DB 1140 GAGCAGAGAACTGTGAGAAAGCAGATGCGCCAGCTGGCGCTGATCCGCGCATGCTGAC 1199
QY 1202 GAGCTGACCAAGAGCGCATCAAGTTCAATCAATGAACGGGCTTATGCGGACCCCATG 1261
DB 1200 GAGCGGAGACAGAGAGATCAAGTTCAATCAATGAATGAGATCTGATGATGACCCCATG 1259
QY 1262 AAGGTGTAAGAAACCGCCAGTCAATGAACATGTGAGAGGAGAGAGAGAGACCTTC 1321
DB 1260 AAGGTGTAAGAGACCGGTACCAATGTGAGAGGAGAGAGAGAGAGACCTTC 1319
QY 1322 CGGAGAGATTCAATGACATCCCAAGAACTTGGCTGATCGCATCATCTCTGAGAGG 1381
DB 1320 CGTGAAGAAATTTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1379
QY 1382 AAGACAGTGGCTGAGTGTCTCTATTAATCACTGATCTAAGAAATGAGAACTATTAAG 1441
DB 1380 NNNACGTCGTCGATGATGTCTCTATTAATCACTGATCTAAGAAATGAGAAATTTCAAG 1439
QY 1442 AGCTGTGAGACGAGCTATGCGCGCGCGGCAAGACAGCAGACAAACAGCAGCAG 1501
DB 1440 AGCTGTGAGAGGAGCTATGCGCGCGGTGCAAGAGCGAGNNNNNNNNNNNNNNNN 1499
QY 1502 CAGCAGACAGACAGCAGCAGCAGCCATGCCCCGACAGCCAGAGAGAGAA 1561
DB 1500 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1559
QY 1562 GATGAGAGAGAGAGAAAGAGGCGAGAGAGAGAGAGAGAGCGGAGGTGAGAAAC 1621
DB 1560 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1619
QY 1622 GACAGAGAAAGACTCTCTCAAGAGAGAGAGAGACACTCAAGGAGAGAGCAACGACGAG 1681
DB 1620 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1679
QY 1682 AAGGAGGCTGTGGCTCAAAAGGCGCAAACTGGCAACAGCCAGGGAAGAGAGCAAGGC 1741
DB 1680 AAGGAGGCGTGGCTCTCAAAAGGCGCAAACTGGCAACAGCCAGGCGCGCAAGGC 1739
QY 1742 CGCATCAACCGCTCAATGAGCTAATGAGCCAAACGAGAGAGGCTATCAACCCGACGAG 1801
DB 1740 CGTATCAACCGCTCAATGAGCTAATGAGCCAAACGAGAGAGAGCAACCCGACGAG 1799
QY 1802 AGCGCGAGCTGGCTCTCAATGAGCTAATGAGCTTCTGCTGAGCAAGAAAGAAAG 1861
DB 1800 AGTTCAAGAGCTGCTCTCAATGAGCTAATGAGAGATTTCTGCTGAGCTAAGAGAGAG 1859
QY 1862 GAAACAGCAAGAAAGTCTCTGGAACAGGCGGCAACTGTGCGCATTCGCGCGAGATG 1921
```



Db	1860	GAGACAGCAAGAAAGAGCCCTCTCTGGAACTGTGGAGAACTGGTCTGAGCCATTGCTCCGCACTG	1919
Qy	1922	GTGGGCTCCAAAGCTGTGTGCGAGTGTAAAGACTTTACTTCACTTCAACAAGAAAGGAG	1981
Db	1920	GTGGGCTCCAAAGACCTGTCTCCAGTGTAAAGACTTTACTTCACTTCAACAAGAAAGGAG	1979
Qy	1982	AACCTCGATGATCTTTGACAGCAACAAGCTGAGATGTAGAGAGAGAGAAAGCGCGG	2041
Db	1980	AACCTGGAAGAAATCTTTGAGCAGCAACAAGCTTAAATGTAGAGAGAGAGAAAGCGCTCGG	2039
Qy	2042	AGGAAGAAAGAAAGCGCGGCGGCGGCGCAAGAGAGAGCTGCAATTCGCGCCCTGTGTG	2101
Db	2040	AGGAAGAAAGAAAGACCCCAAGCTGCGGCGAGGAGAGAGAGAGCTTCCCACTTCCGCT	2099
Qy	2102	GAGATGAGAGATGTAGAGCGGTGCGGCGTGTAGCGGAAATGAGAGAGATGTGTGAGAG	2161
Db	2100	GAGAGCAAGATGTAGAGCATCAGGCGCAAGTGCCTAATGAGAGAGAGTGTGCGAGAG	2159
Qy	2162	GCTGAAGCTTATAGCTCTCTGAGAAATGAGGTGCCAGAGGGGAAATGCAGTGCGCCAGCC	2221
Db	2160	GCAGAAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCT	2219
Qy	2222	ACTGTCAAACAAGCTCAGACACCGAGAGATCCCTCTCTCTCACTGAGGCGCCCAAG	2281
Db	2220	GCT-----CCAGCTCCCAAGCAAAATATCCAG	2249
Qy	2282	GACACAGGCGAGATGGGCCCAAGCCCCCAGCCACTGGGCGCGACGGGCGACCCCCA	2341
Db	2250	GAGC-----	2253
Qy	2342	GGCCCAACCAACCCACACAGAGAGACATCCGGGCGCCCATTTAGCCACCCCGCCTCT	2401
Db	2254	-----	2253
Qy	2402	GAAAGCAACCGAGACCCCTAGGCCCCCAAGCAACCCCATTCGCGCTGTGACCTCTCT	2461
Db	2254	-----	2253
Qy	2462	GTGGTCCCAAGAGAGAGAGAGACCGCAGCAGCGCCCCAGTGGAGAGGGG	2521
Db	2254	-----AGAACTCTAGGCGCGAGAGATGTGATGTGGAAAGCCAGAGAA-----	2296
Qy	2522	GAGAGCAGAAAGCCCCCGCGGCTGTGAGAGCTGGCAGTGGACAAGGGAGGCGAGAG	2581
Db	2297	-----GCCGAGGCGCTGTGAGGAGCCCCCGAG	2324
Qy	2582	CCGCTCAAAGCGAGTGTACAGGAGAACCGAGAGAGGGC---GGCCAAAGGCAAGAC	2638
Db	2325	AGTGTAAAGGTACCAAGAGGAGAGACCGAGGAGAGCCTGTGAACAAACCAAGGAGC	2384
Qy	2639	GCGAGGCGCTGAGGCAACGCGCGAGGGGCGCTCAAGGAGAGAAAGAGAGGGCGG	2698
Db	2385	ACAGAGGCAATTAAACTGTGTCTTAGAGCACTTAAG---GTGAGAGAGGCTGTAGC	2441
Qy	2699	AGCGGCAAGGCAACAATGCAAGAGCTCGGCGCCCCCAGAGACAAGCACTTCAGTCT	2758
Db	2442	AAGCAGAGCTGTGACCAAGGTTTCAGCTCAGGTGCCAACAGACAGTGTCTCCAGTGC	2501
Qy	2759	ACCTGCAGTGACAGAGGTGTGATGAGCGCGAGGGCGCGACAAGAACCGGCTGTGTCC	2818
Db	2502	ACCTGCAGTGCGCATGATGAGGTGTGAGAACCGAAGGAGTGTGACAAGGCGAGGCTGTGTCA	2561
Qy	2819	CCAAAGGCCAGGCTCTCAACCCGACTGGGAGACCCCGGGCCAAAGGCTCAACCCCAAG	2878
Db	2562	CCAAAGGCCAGGCTCTCTCAACCCGCGCTGTGAGATCCCGGGCCAGTACCTCTGCGCCCAAG	2621
Qy	2879	CCAAGTGAAGCTGAGAGAGTGAAGCAGCAGCGGCTGCATCCCGCCATCCAGGTCAAC	2938
Db	2622	CCGCTGGAAGCTGAGAGAGTGAAGCAGCAGCAGCGGCTGCATCCCGCCATCCAGGTCAAC	2681
Qy	2939	AAAGTTCATAGGCCCCCGGAGAGACGAGCTCCCAACCAAGCTCCCGCAAGCCCCCA	2998



[illegible]

Db	6595	AAATGCCGAACACACTCACCACCAACCGTGTGTGGAGAGACCGGCCCTCATCTGCAG	7013
RESULT 4			
LOCUS	BC033087		
DEFINITION	Homo sapiens, similar to nuclear receptor co-repressor 2, clone IMAGE:5016291, mRNA.		
ACCESSION	BC033087		
VERSION	BC033087.1		
KEYWORDS	GI:21619972		
SOURCE	HTC.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Bukacynski, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 2000)		
TITLE	Straussberg, R.		
JOURNAL	Direct Submission		
REMARK	Submitted (25-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
COMMENT	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a> Contact: MGC help desk Email: <a href="mailto:gcgabs-remail.nih.gov">gcgabs-remail.nih.gov</a> Tissue Procurement: DCTD/DTF cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada <a href="mailto:info@cgsc.bc.ca">info@cgsc.bc.ca</a> Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Felli, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Rita Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeary, Steven Ness, Pavan Pandoh, Anna-Lilja Prabhu, Parvaneh Saedi, Jacqueline Schrein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranda Tsai, Natafja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Series: IRAL Plate: 43 Row: 1 Column: 21 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5454073 This clone has the following problem: retained intron. Location/Qualifiers 1..2000 /organism="Homo sapiens" /mol_type="RNA" /db_xref="taxon:9606" /clone="IMAGE:5016291" /tissue_type="lung, small cell carcinoma" /clone_id="NH_MGC_7" /lab_host="DH10B-R" /note="Vector: pOT87"		
ORIGIN			
Query Match	17.2%	Score 1473.6;	DB 4; Length 2000;
Best Local Similarity	96.8%;	Pred. No. 1.8e-265;	
Matches 1503; Conservative	0;	Mismatches 49;	Indels 0; Gaps 0;
1	CATCTCGGGCTCCACAGCTGTGGAGACAGCTGAGGGCCACTGAGCCCGCTACC	60	
272	CATCTCGGGATCCACAGAGCTGTGGAGACAGCTGAGGGCCACTGAGCCCGCTACC	331	
61	GCCTCAGAGCTTCTTACCAAGTGAATGCGCCGACGACAGCAAGTCTGGGCTCT	120	
332	GCCTCAGAGCTTCTTACCAAGTGAATGCGCCGACGACAGCAAGTCTGGGCTCT	391	
121	GAGATACAGACCACTCCGCGACATATGCTCCACCTGTGCGCGGGCTCATATCA	180	

Db 392 GAGATACAGACCACTCCCGCACTATGCTCCCACTGCGCCGCTCATCATCA 451  
 Qy 181 GCCCAGCGGAGGCGCTCCCTGCTGTGATTCAGCCCGGAATGAACGGTCCCA 240  
 Db 452 GCCCAGCGGAGGCGCTCCCTGCTGTGATTCAGCCCGGAATGAACGGTCCCA 511  
 Qy 241 GAGCTTCACCTGCGGAGAGTCCCATATACCTGCGGAGTGGGAATGACAGAT 300  
 Db 512 GAGCTTCACCTGCGGAGAGTCCCATATACCTGCGGAGTGGGAATGACAGAT 571  
 Qy 301 GAGCTTCATTTGAAAGAGGCGCTGCGCTGATGAGTCTGCTGACCCCTGCTGCA 360  
 Db 572 GAGCTTCATTTGAAAGAGGCGCTGCGCTGATGAGTCTGCTGACCCCTGCTGCA 631  
 Qy 361 GTACCCCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
 Db 632 GTACCCCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 691  
 Qy 421 CTTGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
 Db 692 CTTGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 751  
 Qy 481 GAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
 Db 752 GAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 811  
 Qy 541 CCGAGAGATCACTGATGATGATGATGATGATGATGATGATGATGATGAT 600  
 Db 812 CCGAGAGATCACTGATGATGATGATGATGATGATGATGATGATGATGAT 871  
 Qy 601 GAGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 Db 872 GAGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 931  
 Qy 661 GTGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
 Db 932 GTGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 991  
 Qy 721 TGACATCGGATCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 780  
 Db 992 TGACATCGGATCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 1051  
 Qy 781 CTCGACACCCGAGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 Db 1052 CTCGACACCCGAGATGATGATGATGATGATGATGATGATGATGATGAT 1111  
 Qy 841 AATCTTGTACTTGAAGAGGAGATGATGATGATGATGATGATGATGATGAT 900  
 Db 1112 AATCTTGTACTTGAAGAGGAGATGATGATGATGATGATGATGATGATGAT 1171  
 Qy 901 GGGCTATGACAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 Db 1172 GGGCTATGACAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1231  
 Qy 961 GCGCGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
 Db 1232 CCGGCGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1291  
 Qy 1021 CCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
 Db 1292 CCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1351  
 Qy 1081 GGTGTCATGTCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
 Db 1352 GGTGTCATGTCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1411  
 Qy 1141 AAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
 Db 1412 AAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1471  
 Qy 1201 CGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
 Db 1472 CGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1531

Qy 1261 GAAGTGTACAAAGACCGGAGGATGATGATGATGATGATGATGATGATGAT 1320  
 Db 1532 GAAGTGTACAAAGACCGGAGGATGATGATGATGATGATGATGATGATGAT 1591  
 Qy 1321 CCGGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
 Db 1592 CCGGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1651  
 Qy 1381 GAAGACAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440  
 Db 1652 GAAGACAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1711  
 Qy 1441 GAGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500  
 Db 1712 GAGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1771  
 Qy 1501 GCAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1552  
 Db 1772 GCAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1823

RESULT 5  
 BC020427 1362 bp mRNA linear HTC 03-JAN-2002  
 LOCUS  
 DEFINITION Homo sapiens, clone IMAGE:4179307, mRNA.  
 ACCESSION BC020427  
 VERSION BC020427.1 GI:18043012  
 KEYWORDS  
 SOURCE HTC.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

REFERENCE  
 AUTHORS Strausberg, R.  
 TITLE Direct Submision  
 JOURNAL Submitted (03-JAN-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,  
 USA

REMARK  
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: David N. Louis, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 Sequencing Center  
 Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [villalombcm.tmc.edu](mailto:villalombcm.tmc.edu)  
 Villalon, D.K., Luna, R.A., Hale, S.M., Huiyk, S., Lu, X., Garcia,  
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,  
 Muzny, D.M., Guarnante, P., Yoon, V., Kowis, C., Martin, R.,  
 Lawrence, S., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
 Series: IRAX Plate: 12 Row: 1 Column: 23  
 This clone has the following problem: no polyA-tail.

FEATURES  
 Location/Qualifiers

source

1..1362  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4179307"  
 /tissue\_type="Brain, anaplastic oligodendroglioma with  
 1p19q loss"  
 /clone\_lib="NCI CGAP\_Brn67"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

ORIGIN







QY 6903 AGCCTGAATACATATATACAGCCAGCCTGGACGAGATCTTCAATATGCCCGCATCACCG 6962  
DB 361 AGCCTGAATACATATATACAGCCAGCCTGGACGAGATCTTCAATATGCCCGCATCACCG 420  
QY 6963 GAAACGGCTTATATGACCTTATAGAACCGGCTGACGAAATATGCCAGACCAATATG 7022  
DB 421 GAAACGGCTTATATGACCTTATAGAACCGGCTGACGAAATATGCCAGACCAATATG 480  
QY 7023 GAGCTGAGGCGCATATATATAGAAAGCACTCATGGTAAATATAGACAGTGGAAAGTCCC 7082  
DB 461 GAGCTGAGGCGCATATATATAGAAAGCACTCATGGTAAATATAGACAGTGGAAAGTCCC 540  
QY 7083 CGCGCTCAGCGCCCAATGCTTTTAAACCTCTGATATGCGAGTCCAGCTGCGCGTGTCTA 7142  
DB 541 CGCGCTCAGCGCCCAATGCTTTTAAACCTCTGATATGCGAGTCCAGCTGCGCGTGTCTA 600  
QY 7143 TGCCCATTAACCGCTGCTGACGAGAGTACCACTCACTCTGCGCAGTGGCGCGC 7202  
DB 601 TGCCCATTAACCGCTGCTGACGAGAGTACCACTCACTCTGCGCAGTGGCGCGC 660  
QY 7203 GGAAGGCAAGGCTCTGACGAGACCGAGCGGAAAGCAAGTCCCGCGCGCGCGCGC 7262  
DB 661 GGAAGGCAAGGCTCTGACGAGACCGAGCGGAAAGCAAGTCCCGCGCGCGCGCGC 720  
QY 7263 TGCGATCTGAGGAGCGGCGCACCTCTGTCTCTCAATGACACTGAGAGGAGACTGCAAC 7322  
DB 721 TGCGATCTGAGGAGCGGCGCACCTCTGTCTCTCAATGACACTGAGAGGAGACTGCAAC 780  
QY 7323 GCGGAGCGCGCTGACCAACCGCGGTGAGAGAGACAGCGCTCTGCGCAGTTCACAGC 7382  
DB 781 GCGGAGCGCGCTGACCAACCGCGGTGAGAGAGACAGCGCTCTGCGCAGTTCACAGC 840  
QY 7383 CATTCGCCCTACACCGCCCTGATCATGC--GGCTGACAGCGGAGT--GTATAGCTTCCCCAC 7439  
DB 841 CATTCGCCCTACACCGCCCTGATCATGCNNGTGACGCGGAGTGTGATGTTTCCCA 900  
QY 7440 CCCCAACCGGCGCTCCCGCGGAGCGG 7466  
DB 901 CCCCAACCGGCGCTCCCGCGGAGG 927

RESULT 8  
B0557144 846 bp mRNA linear EST 16-SEP-2002  
LOCUS AGSCOURT\_10253238 NIH MGC 109 Homo sapiens cDNA clone  
DEFINITION IMAGE:6584961 5', mRNA sequence.  
ACCESSION B0557144 GI:22907440  
VERSION B0557144.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 846)  
NIH-MGC <http://mhc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: Agencourt Bioscience Corporation (LLNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1CM2796 row: c column: 09  
High quality sequence stop: 719.  
Location/Qualifiers  
1..846  
Source /organism="Homo sapiens"

/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6584961"  
/lab\_type="teratocarcinoma, cell line"  
/issue\_type="DH10B (phage-resistant)"  
/lib host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_109"  
/note="Organ: ovary; Vector: pOTB1; Site 1: EcoRI; Site 2:  
XhoI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
SuperScript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

Query Match 9.8%; Score 841.8; DB 5; Length 846;  
Best Local Similarity 99.8%; Pred. No. 4.8e-147;  
Matches 843; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6730 GCTGCTGACCGGAGTGGGGAACAGACGAGCCGACGAGATGGGCTTCAAGTCTCCAGG 6789  
DB 1 GCTGCTGACCGGAGTGGGGAACAGACGAGCCGACGAGATGGGCTTCAAGTCTCCAGG 60  
QY 6790 CAACACGAGCGACCGCGCTTCTTACGACAGTGAACGAGACCACTCCGCAATGT 6849  
DB 61 CAACACGAGCGACCGCGCTTCTTACGACAGTGAACGAGACCACTCCGCAATGT 120  
QY 6850 CAAGTCCAAAGAGCAAGATCAACAAGAACTGAACCCACAAACGGAATGAGCTGA 6909  
DB 121 CAAGTCCAAAGAGCAAGATCAACAAGAACTGAACCCACAAACGGAATGAGCTGA 180  
QY 6910 ATACATATACAGCAGCCTGAGAGAGATCTTCAATATGCGCCGATCACCGAAACGG 6969  
DB 181 ATACATATACAGCAGCCTGAGAGAGATCTTCAATATGCGCCGATCACCGAAACGG 240  
QY 6970 CCTTATGACCTTATAGAACCGGCGGTGACAGAAATGTCAGACCAATATGGGCTGGA 7029  
DB 241 CCTTATGACCTTATAGAACCGGCGGTGACAGAAATGTCAGACCAATATGGGCTGGA 300  
QY 7030 GAGCATATTTAGAAAGGCACTCATAGGTAAATATGACAGTGGAGAGTCCCGCGCT 7089  
DB 301 GAGCATATTTAGAAAGGCACTCATAGGTAAATATGACAGTGGAGAGTCCCGCGCT 360  
QY 7090 CAGCGCCAAATGCTTTTAAACCTCTGAAATGCAAGTCCAGCTGCCGCTGCTATGCCAT 7149  
DB 361 CAGCGCCAAATGCTTTTAAACCTCTGAAATGCAAGTCCAGCTGCCGCTGCTATGCCAT 420  
QY 7150 AACCGCTGCTGACGAGCGAGTACCACTCACTCTGCGCAGTGGCGCGGAGAGGC 7209  
DB 421 AACCGCTGCTGACGAGCGAGTACCACTCACTCTGCGCAGTGGCGCGGAGAGGC 480  
QY 7210 CAAGTCTCTGAGAGACCGAGCGGAAAGGCAAGTCCCGCGCGCGGCTGAGCATC 7269  
DB 481 CAAGTCTCTGAGAGACCGAGCGGAAAGGCAAGTCCCGCGCGCGGCTGAGCATC 540  
QY 7270 TGAGGAGCGGACCACTCTGTCTCTCTCACTGACCTGAGAGGAGATGCAACCGCGGAC 7329  
DB 541 TGAGGAGCGGACCACTCTGTCTCTCTCACTGACCTGAGAGGAGATGCAACCGCGGAC 600  
QY 7330 GCGGCTGACCAACCGCGGTGAGAGAGACAGCCCTGCTGCGCAGTTCAGAGCATTTCCC 7389  
DB 601 GCGGCTGACCAACCGCGGTGAGAGAGACAGCCCTGCTGCGCAGTTCAGAGCATTTCCC 660  
QY 7390 CTACAAACCCCTGATCATAGCGGCTGACGCGGAGTGTATGCTTCCCAACCCCAACGCGG 7449  
DB 661 CTACAAACCCCTGATCATAGCGGCTGACGCGGAGTGTATGCTTCCCAACCCCAACGCGG 720  
QY 7450 CTTCCCGCGGAGCGGCGCTCTGCTGAGCCGACCAACGCTGAGAGAGAGAGCCCA 7509  
DB 721 CTTCCCGCGGAGCGGCGCTCTGCTGAGCCGACCAACGCTGAGAGAGAGAGCCCA 780  
QY 7510 GCGACTGCTTGTGCTGCGATGAGAGCACTCTCCGACAGCGAGTGACTCAAGAACGGGCG 7569

Db 781 GGCACGCTCTGCTGCAGTAGAGACACTCTCCGACAGAGATGACTCAGAAACGACCG 840  
QY 7570 GGGGG 7574  
Db 841 GGGGG 845

RESULT 9  
LOCUS BM553310 1126 bp mRNA linear EST 20-FEB-2002  
DEFINITION AGENCOURT 6572642 NIH\_MGC\_41 Homo sapiens cDNA clone IMAGE:5467163  
5', mRNA sequence.  
ACCESSION BM553310  
VERSION BM553310.1 GI:18791930  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1126)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-r@mail.nih.gov  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM1968 row: 1 column: 12  
High quality sequence stop: 674.  
Location/Qualifiers  
1. 1126  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5467163"  
/tissue\_type="amelanotic melanoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_41"  
/note="Organ: skin; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene),  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

ORIGIN

Query Match 9.5%; Score 815.6; DB 3; Length 1126;  
Best Local Similarity 99.5%; Pred. No. 4e-142;  
Matches 818; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4369 GGTCAAGTACGACACGGGCGCTCCACACTGCTCCAAAGACAGAGTACGCTCCCT 4428  
Db 1 GGTCAAGTACGACACGGGCGCTCCACACTGCTCCAAAGACAGAGTACGCTCCCT 60

QY 4429 CATCGGACGCCCCCGCGCGAGCTTCCACCCGTGCACCCGCTGATGTATGCGCCAGC 4488  
Db 61 CATCGGACGCCCCCGCGCGAGCTTCCACCCGTGCACCCGCTGATGTATGCGCCAGC 120

QY 4489 CCGGGAGCTGGAACGCGCTCTACAGAGAGAGCTGAAGAGCGGCGCAGGCGCCAG 4548  
Db 121 CCGGGAGCTGGAACGCGCTCTACAGAGAGAGCTGAAGAGCGGCGCAGGCGCCAG 180

QY 4549 CAGCTCGGGGGGCTCCATTGCGCGCGCGCGCGCTCATTTGCTGAGCTGGTAAAGCC 4608

Db 181 CAGCTCGGGGGGCTCCATTGCGCGCGCGCGCGCTCATTTGCTGAGCTGGGCAAGCC 240  
QY 4609 GCGGACAGAGCCCCCTGACCTATAGAGACACAGGGGACCTTTGCGGACACTCCACG 4668  
Db 241 GCGGACAGAGCCCCCTGACCTATAGAGACACAGGGGACCTTTGCGGACACTCCACG 300

QY 4669 AGTTTCGCCCCGTGACCATAGCGGAGGCCACGCGCGCTGACAGAGAGGAGCTTTGCTC 4728  
Db 301 AGTTTCGCCCCGTGACCATAGCGGAGGCCACGCGCGCTGACAGAGAGGAGCTTTGCTC 360

QY 4729 CAGCAAGGATCCGAGAGACCGAAAGCTGACGCTCTGATGATGATGCCAAGTCCCC 4788  
Db 361 CAGCAAGGATCCGAGAGACCGAAAGCTGACGCTCTGATGATGATGCCAAGTCCCC 420

QY 4789 GCAACAGACCGGCGCGGACACACCCACATCTGCGCTATGACACCTGCTTG 4848  
Db 421 GCAACAGACCGGCGCGGACACACCCACATCTGCGCTATGACACCTGCTTG 480

QY 4849 GGGCGTGAATGCGCTGAGACTGTATGACAGCACATCCCCCTTGACCTTGCACCTTC 4908  
Db 481 GGGCGTGAATGCGCTGAGACTGTATGACAGCACATCCCCCTTGACCTTGCACCTTC 540

QY 4909 CATACCCGCGGCATCTCTTGAAGAGAGCGCGCTGCTACTACCTGCCGACACTGGC 4968  
Db 541 CATACCCGCGGCATCTCTTGAAGAGAGCGCGCTGCTACTACCTGCCGACACTGGC 600

QY 4969 CCCCAACCCCACTACCCGACCTGATCCACCTACCTACCTACCTACCTACCTACCTAC 5028  
Db 601 CCCCAACCCCACTACCCGACCTGATCCACCTACCTACCTACCTACCTACCTACCTAC 660

QY 5029 GCGCGGCTGGAAGAACCGGACAGACATCATCATGACTACATCTGACAGAGATGCA 5088  
Db 661 GCGCGGCTGGAAGAACCGGACAGACATCATCATGACTACATCTGACAGAGATGCA 720

QY 5089 CCAACAACAGCGCCACCGGCATGCGCCAGCGAGCTGATGAGGGGCTCTCGCCCG 5148  
Db 721 CCAACAACAGCGCCACCGGCATGCGCCAGCGAGCTGATGAGGGGCTCTCGCCCG 780

QY 5149 CGAGTCTCTGCTGCGACCTCACTAGCGTGGGTCGCCGAGG 5190  
Db 781 CGAGTCTCTGCTGCGACCTCACTAGCGTGGGTCGCCGAGG 822

RESULT 10  
LOCUS BQ690869 984 bp mRNA linear EST 15-JUL-2002  
DEFINITION AGENCOURT 8034173 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6208241  
5', mRNA sequence.  
ACCESSION BQ690869  
VERSION BQ690869.1 GI:21816185  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 984)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2365 row: 3 column: 18  
High quality sequence stop: 671.  
Location/Qualifiers  
1. 984

```

"organism":"Homo sapiens"
/mol_type:"mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6208241"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1b="NIH MGc 110"
/notes="Organ: pancreas; Vector: pOMB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGc library."

```

Query Match	9.3%	Score 798.2;	DB 5;	Length 984;
Best Local Similarity	94.7%;	Pred. No. 7.3e-139;		
Matches 892;	Conservative	0;	Mismatches 39;	Indels 11;
				Gaps 6;

4001 GCCATCCCGCCGGAGCGACACAGCCCCACCACTCAAAGAGCAGCACCATTCGGCGG 4066

4061 TCCATCACACAAGGATCCCTCGGTTCTACGTGAGGCACAGAGACTACCTGCGTGG 4120  
61 TCCATCACACAAGGATCCCTCGGTTCTACGTGAGGCACAGAGACTACCTGCGTGG 120

4121 GAGGCCAAGCTCCTAAAGCGGGAGGGCAGCGCTCCGCCCCCAACCGCCTCAGCGGAGCTGG 4180  
121 GAGGCCAAGCTCCTAAAGCGGGAGGGCAGCGCTCCGCCCCCAACCGCCTCAGCGGAGCTGG 180

4181 ACCGAGGCTACAAGACGAGGCCCTGGGCCCCCTGAAGCTAAGCCGGCCCATGAGGCG 4240

181 ACCGAGGCTACAAGACGAGGCCCTGGGCCCCCTGAAGCTGAAGCGGCGCCATGAGGCG 240

4241 CTGTGGCCACGGTGAAGAGCGGGCCGCTTCATCATGAAATCCGCGGAGAGCTG 4300  
241 CTGTGGCCACGGTGAAGAGCGGGCCGCTTCATCATGAAATCCGCGGAGAGCTG 300

4301 CGGCACACGCCGAGTGCCTTGGCCCGCGCGCTCAAGAGGGCTCCATCACGAG 4306  
301 CGGCACACGCCGAGTGCCTTGGCCCGCGCGCTCAAGAGGGCTCCATCACGAG 360

4361 GGCACCCCGCTCAAGTACGACACCGCGCTCCACCTGGCTCCAAAAGCAGACGTA 4420  
361 GGCACCCCGCTCAAGTACGACACCGCGCTCCACCACTGGCTCCAAAAGCAGACGTA 420

4421 CGCTCCCTCATCGGCAGCCCGCGCCGGACGTTCCACCCCGTGCATGTGATG 4480  
|||  
421 CGCTCCCTCATCGGCAGCCCGCGCCGGACGTTCCACCCCGTGCATGTGATG 480

4481 GCCGACGCCCGGCACTGGACGTGCTCTGACGAGAGAGCTGAAGAGCGCCGCAAGG 4510  
481 GCCGACGCCCGGCACTGGACGTGCTCTGACGAGAGAGCTGAAGAGCGCCGCAAGG 540

4541 ACCGCCAAG-CAGCTCGGGGGCTCATTGCGCGCGGCCCGGTCAATTGCTGAGCT 4539  
541 ACCGCCAAGCTCGGGGGCTCATTGCGCGCGGCCCGGTCAATTGCTGAGCT 600

4600 GGGTAAGCCGCGCAGAGCCCTCTGACTATGAGGACACCGGGCACCCTTGGCCGGCCA 4655  
601 GGGTAAGCCGCGCAGAGCCCTCTGACTATGAGGACACCGGGCACCCTTGGCCGGCCA 660

4660 CCTCCACGAGGTTGCGCCGTGACATGCGGAGGCCACGCGCGCTGCAGGAGGGCAG 4719  
661 CCTCCACGAGGTTGCGCCGTGACACGCGGAGGCCACGCGCGCTGCAGGAGGGCAG 720

4720 CCTTTCGTCACGACAGGCAATCCCA-GGACCGAAAGCTGACGTCGACGCTCGAGATCG 4778  
721 CCTTTCGTCACGACAGGCAATCCCAAGCTGACGTCGACGCTCGAGATCG 780

QY	4779	CCAAAGTCCCGCAGAGACCGTGGCCGAGACACACCAACCCG--ATCTGGCCCTATGA	4836
Db	781	CCAAAGTCCCGCAGACCGTGGCCGAGACACCAACCCCACTCTGGCCCTATGA	840
QY	4837	GCACCTG----CTTCGGGGCGTGATGTGGCGTGACCTGTATCGAGGCAC--ATCCCGCT	4890
Db	841	GCACCTTGGCTTCCGGGGCGTGAAATGGGCGTGAATCTGTATCCCAACCACCATCCCCCTG	900
QY	4891	GGCGCTTGACCCCACTTCATA--CCCGGGCATTCCTCTGG	4931
Db	901	GGCGCTTGAAACCCACTTCATACCCCGGGCATTCCTCTGG	942

LOCUS	BM423558	1051 bp	mRNA	linear	EST 29-JAN-2002
DEFINITION	AGENCOURT_6399411 NIH_MGC_41 Homo sapiens cDNA IMAGE:5516745				
ACCESSION	BM423558	5', mRNA sequence.			
VERSION	BM423558.1	GI:18391770			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

REFERENCE	1 (Pages 1 to 1051)
AUTHORS	NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strauberg, Ph.D.

CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.llnl.gov>  
Plate: LNCM2017 row: n column: 10  
High quality sequence stop: 667.  
Location/Qualifiers

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5516745"
/issue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 41"
/note="Organ: skin; Vector: pOTB7; site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5 adaptor:
GGGACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

```

Query Match	9.3%	Score 796.4	DB 3	Length 1051
Best Local Similarity	98.0%	Pred. No. 1.6e+18		
Matches 817	0	Mismatches 16	Indels 1	Gaps 1

4369 GGTCAAAGTACGACACCGGCGCGTCCACT-GGCTCCAAAAAAGCAGATAGCTCCC 4427  
 1 GCTCAAGTACGACTCCGGCGCGTCCACCACTGGGCTCCAAAAAAGCAGATAGCTCCC 60

Db 61 TCATCGCAGCCCCCGCCGAGCCTTCCCAACCCGTGACCCCGCTTGATGTGATGACCCGACG 120

4488 CCGGGCACTGGAACTGCTCTACGAGGAGAGCCTGAAGAGCCGGCCAGGACCGCA 4547

|||||  
Db 121 CCGGGGCACTGGAACGCTGCTACGAGAGAGCCGGAAGACCCGGCCAGGGACCGCCA 180  
|||  
QY 4548 GCAGCTCGGGGGGCTCTCATTTGCGCGGGGCCCGGCTCATTTGGCTGAGCTGGTAAAC 4607  
|||  
Db 181 GCAGCTCGGGGGGCTCTCATTTGCGCGGGGCCCGGCTCATTTGGCTGAGCTGGGCAAGC 240  
|||  
QY 4608 CGGGGAGAGCCCCCTGACCTATGAGAGCAACGAGGGGACCCCTTTGGCGGGCACCCTCCAC 4667  
|||  
Db 241 CGGGGAGAGCCCCCTGACCTATGAGAGCAACGAGGGGACCCCTTTGGCGGGCACCCTCCAC 300  
|||  
QY 4668 GAGGTTGCGCCCTGACCATGCGGGAGCCCAACCGCGGCTGACGAGAGGAGGCTTTTCT 4727  
|||  
Db 301 GAGGTTGCGCCCTGACCATGCGGGAGCCCAACCGCGGCTGACGAGAGGAGGCTTTTCT 360  
|||  
QY 4728 CCGAGAGGATCTCCAGAGACCGAAGAGTGAAGCTGAGAGGCTCTGAGATGCGCAATCCC 4787  
|||  
Db 361 CCGAGAGGATCTCCAGAGACCGAAGAGTGAAGCTGAGAGGCTCTGAGATGCGCAATCCC 420  
|||  
QY 4788 CGCAGAGACCGGCGCCGAGACCAACCCCACTCTGCGCCCTATGAGACCTGCTTC 4847  
|||  
Db 421 CGCAGAGACCGGCGCCGAGACCAACCCCACTCTGCGCCCTATGAGACCTGCTTC 480  
|||  
QY 4848 GGGGCGTGAAGTGGCTGAGACCTGATGACAGCCACATCCCTGCGCTTGCACCCCACT 4907  
|||  
Db 481 GGGGCGTGAAGTGGCTGAGACCTGATGACAGCCACATCCCTGCGCTTGCACCCCACT 540  
|||  
QY 4908 CCATACCCCGCGGACATCCCTCTGAGAGCGACGCGCTGCTAATCTGCGCCGACACTGG 4967  
|||  
Db 541 CCATACCCCGCGGACATCCCTCTGAGAGCGACGCGCTGCTAATCTGCGCCGACACTGG 600  
|||  
QY 4968 CCGCCCAACCCCACTACCCGCGACCTGATACCACTGATCCGCGGCTACCCGAGCA 5027  
|||  
Db 601 CCGCCCAACCCCACTACCCGCGACCTGATACCACTGATCCGCGGCTACCCGAGCA 660  
|||  
QY 5028 CCGCGCGCTGAGAGACCGGAGACCATCATCATGATGATGAGAGGCTCTCGCCCC 5087  
|||  
Db 661 CCGCGCGCTGAGAGACCGGAGACCATCATCATGATGATGAGAGGCTCTCGCCCC 720  
|||  
QY 5088 ACCACAAACCGGACCGGACCATGAGCTGATGAGAGGCTCTCGCCCC 5147  
|||  
Db 721 ACCACAAACCGGACCGGACCATGAGCTGATGAGAGGCTCTCGCCCC 780  
|||  
QY 5148 GCGAGCTCTGCGGACCTCACTACGCTGCGGGTCCCGAGGACATCATGAGCC 5201  
|||  
Db 781 GCGAGCTCTGCGGACCTCACTACGCTGCGGGTCCCGAGGAGGACCAACC 834  
|||  
RESULT 12  
BM802749 1159 bp mRNA 1linear EST 05-MAR-2002  
LOCUS BM802749  
DEFINITION AGENCOURT\_6459836 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:5575037  
5', mRNA sequence.  
ACCESSION BM802749  
VERSION BM802749.1 GI:19119572  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 1159)  
NIH-MGC <http://mgc.nci.nih.gov/>,  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph. D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
CDNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:

<http://image.lnl.gov>  
Plate: L1AM12324 row: k column: 06  
High quality sequence stop: 652.  
Location/Qualifiers  
1. 1159  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5575037"  
/issue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; Oligo-dT primed.  
Average insert size 2.5 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."  
ORIGIN  
Query Match 9.3%; Score 796.2; DB 3; Length 1159;  
Best Local Similarity 94.4%; Pred. No. 1.8e-138;  
Matches 869; Conservative 0; Mismatches 48; Indels 4; Gaps 4;  
QY 4621 CCTGACCTATGAGAGACACGAGGACACCTTTGCGGCACTTCCACGAGTTGCGCCGT 4680  
|||  
Db 1 CCGACCTATGAGAGACACGAGGACACCTTTGCGGCACTTCCACGAGTTGCGCCGT 60  
|||  
QY 4681 GACCATGCGGAGACCCACGCGCGCTGACAGAGGACACCTTTCTCCAGCAAGGCATC 4740  
|||  
Db 61 GACCATGCGGAGACCCACGCGCGCTGACAGAGGACACCTTTCTCCAGCAAGGCATC 120  
|||  
QY 4741 CCAGAGCGAAGAGCTAGCGTGCAGCGCTCGTAGATGCGCAAGTCCCGACAGACCGT 4800  
|||  
Db 121 CCAAGACCGAAGAGCTAGCGTGCAGCGCTCGTAGATGCGCAAGTCCCGACAGACCGT 180  
|||  
QY 4801 GCCCGAGACCAACCCACACCCACTGCGCTATGAGACCTTGCTTGCGGGCTGAGTG 4860  
|||  
Db 181 GCCCGAGACCAACCCACACCCACTGCGCTATGAGACCTTGCTTGCGGGCTGAGTG 240  
|||  
QY 4861 CGTGAACCTGATGCGAGGACCATACCCCTGCGCTTGACACCCACTCCATACCCCGCG 4920  
|||  
Db 241 CGTGAACCTGATGCGAGGACCATACCCCTGCGCTTGACACCCACTCCATACCCCGCG 300  
|||  
QY 4921 CATCCCTGAGAGCGAGCGGCTGCTACTACCTGAGCCCGACACTGAGCCCGCAACCCAC 4980  
|||  
Db 301 CATCCCTGAGAGCGAGCGGCTGCTACTACCTGAGCCCGACACTGAGCCCGCAACCCAC 360  
|||  
QY 4981 CTACCGGACCTGTATACCAACCTTACTCTACCGGCTTACCCCGACAGCGGCGCTGGA 5040  
|||  
Db 361 CTACCGGACCTGTATACCAACCTTACTCTACCGGCTTACCCCGACAGCGGCGCTGGA 420  
|||  
QY 5041 GAAACCGGACCATATATGATGATCACTACCTGCGAGAGATGACCAACACCGCG 5100  
|||  
Db 421 GAAACCGGACCATATATGATGATCACTACCTGCGAGAGATGACCAACACCGCG 480  
|||  
QY 5101 CACCGGACATGGCCAGCGAGCTGATGCTGAGGGGCTCTGCGCCCGGAGAGTCTCGCT 5160  
|||  
Db 481 CACCGGACATGGCCAGCGAGCTGATGCTGAGGGGCTCTGCGCCCGGAGAGTCTCGCT 540  
|||  
QY 5161 GGCATCTCACTACGCTGCGGAGTCCCGAGGACATCATGACTGTGCCAAGTGCACACT 5220  
|||  
Db 541 GGCATCTCACTACGCTGCGGAGTCCCGAGGACATCATGACTGTGCCAAGTGCACACT 600  
|||  
QY 5221 GCTGTGCTGTCGCCCCCGACACACGAGACCCCAAGCCCATGAGACCGCTTGCCTTA 5280  
|||  
Db 601 GCTGTGCTGTCGCCCCCGACACACGAGACCCCAAGCCCATGAGACCGCTTGCCTTA 660  
|||  
QY 5281 CCTCCCAACGCGGCGCCAG-CCCTCAGACGCGGACAGAGCTCCCACTCTCCAG 5339  
|||  
Db 661 CTTCCCAACGCGGCGCCAGCCCTTTCAGAGCGGACAGAGCTCCCACTCTCCAG 720  
|||  
QY 5340 GAGGTCACACACTTGACAAAACCAACCAACGCTGCTGTCGAGCGGAGGAGACC 5399  
|||



Db 721 GAGGTCCACACACTTGACAAACCAACCAACGCTCTCTCCGACGGAGCGAAACC 780  
QY 5400 GGGATCGAGAG-CGGACACGGGATC-GGAGCGGGGAAAGTCATCTTCACCTCCACC-A 5456  
Db 781 GGAATCGAGAGACGGGACCGGGATCGGGGACGGGAAAGTCCTTCTTCACCTCCACCA 840  
QY 5457 CGACGTGAGACGACGACCTCATCTGAGACCTGTGTCAGAGACGACGGCAGCAGCG 5516  
Db 841 CAACGGTGGGACCCCGACCTCTGTGGGAAACCTGTGAACCGAAGAAATATACCGCAGCA 900  
QY 5517 GCAGCAGCGCGGGGGTGGGG 5537  
Db 901 AGGGAACACACAGAGGGGG 921

RESULT 13  
BM477568 1057 bp mRNA linear EST 05-FEB-2002  
LOCUS AGENCOURT\_6484946 NIH\_MGC\_85 Homo sapiens cDNA clone IMAGE:5554420  
DEFINITION 5', mRNA sequence.  
ACCESSION BM477568  
VERSION BM477568.1 GI:18526610  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
1 (bases 1 to 1057)  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Lou Staudt  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LMNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LMNL at:  
http://image.llnl.gov  
Plate: L1AM12272 row: p column: 05  
High quality sequence stop: 674.  
Location/Qualifiers  
1..1057  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5554420"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 85"  
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.867 Kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 9.3%; Score 795.2; DB 3; Length 1057;  
Best Local Similarity 93.5%; Pred. No. 2.7e-138;  
Matches 906; Conservative 0; Mismatches 29; Indels 34; Gaps 6;

QY 2786 GCCGAGGGCGGCGAAGAAACCGGCTGCTGCCCAAGGCCAGGCTCCTCAACCCGAGCT 2845  
Db 1 GCCGAGGGCGGCGAAGAAACCGGCTGCTGCCCAAGGCCAGGCTCCTCAACCCGAGCT 60  
QY 2846 GCGGACCCCGGGGCAATGCTCTACCCGAGGACCACTGGAAGCAAGCTGAAGCAG 2905  
Db 61 GCGGACCCCGGGGCAATGCTCTACCCGAGGACCACTGGAAGCAAGCTGAAGCAG 120  
QY 2906 CGAGGGCTGCTATCCCTCCATCCAGGTCAACCAAGTCAAGAGCCCTCCCGGAGAGC 2965

Db 121 CGAGGCGTGCATCCCTCCCATCCAGGTCAACCAAGTCTAGACCCCTCCGGAGAGC 180  
QY 2966 GCAGCTCCCAACCAAGCAGCTCCCTCAGCCCTCCACCGCACCAAACTTCAGCCGAG 3025  
Db 181 GCAGTCTCCCAACCAAGCAGCTCCCTCAGCCCTCCACCGCACCAAACTTCAGCCGAG 240  
QY 3026 AGCGAGCCCTCAGACGCTTCGAGCAGAGCCCTCCGGGCAAGAGCAGAGCCCGGCACTC 3085  
Db 241 AGCGAGCCCTCAGACGCTTCGAGCAGAGCCCTCCGGGCAAGAGCAGAGCCCGGCACTC 300  
QY 3086 CCGGCGGCAAGAGG-----GCTTCGAGCCGAGGCCAG 3121  
Db 301 CCGGCGGCAAGAGGAGAGAGAGGCTGTGTTCTTCCAGCTTCGAGCCGAGGCCAG 360  
QY 3122 AAGCTGCTGGGAGACCCCTTGTGCTGACTTCGCGGCTGCTCCGCTCCCGTCCCGCCGCT 3181  
Db 361 AAGCTGCTGGGAGACCCCTTGTGCTGACTTCGCGGCTGCTCCGCTCCCGTCCCGCCGCT 420  
QY 3182 GAGGTATCAAGGCTCTCCGCAATGCTCCGAGACCCCTCAAGCTTCTCTACGCTCACCT 3241  
Db 421 GAGGTATCAAGGCTCTCCGCAATGCTCCGAGACCCCTCAAGCTTCTCTACGCTCACCT 480  
QY 3242 GGTCAACCACTGCGGCTGAGCTCAATGACATGCGCGGCGCGTCTGCGGCGGCAACC 3301  
Db 481 GGTCAACCACTGCGGCTGAGCTCAATGACATGCGCGGCGCGTCTGCGGCGGCAACC 540  
QY 3302 ACCATCTCAACCGCGCTCCCTCATCTCTCTGCAAGACCCGAGGCTCTGAGAGG 3361  
Db 541 ACCATCTCAACCGCGCTCCCTCATCTCTCTGCAAGACCCGAGGCTCTGAGAGG 600  
QY 3362 CAATAGGTGCTATCTCCCAAGAAATGTGGTCAAGTCAAGTCCCTACTCAGAGCAT 3421  
Db 601 CAATAGGTGCTATCTCCCAAGAAATGTGGTCAAGTCAAGTCCCTACTCAGAGCAT 660  
QY 3422 GCCAAGGCGCGGTTGGGCTGTCACCAATGGGGTCCCGCCATGAGACCCCAAAAG 3481  
Db 661 GCCAAGGCGCGGTTGGGCTGTCACCAATGGGGTCCCGCCATGAGACCCCAAAAG 720  
QY 3482 CTGGCACCCTTCAGCGAGTGAAGAGAGCAGTGTCTCCACGAGGCGCAGGCTTGAGGCA 3541  
Db 721 CTGGCACCCTTCAGCGAGTGAAGAGAGCAGTGTCTCCACGAGGCGCAGGCTTGAGGCA 780  
QY 3542 CCGGAGAGCTTGGGGGTGCCCAAGCCAGAGGCGTCTGCTGAGAGGACAGCTTG 3601  
Db 781 CCGGAGAGCTTGGGGGTGCCCAAGCCAGAGGCGTCTGCTGAGAGGACAGCTTG 838  
QY 3602 GGTCTAGTTCCGGGCGGAGGATCAACCAAGGCAATGCC---AGACAAGGTTGCCCTCG 3658  
Db 839 GGTCTAGTTCCGGGCGGAGGATCAACCAAGGCAATGCCCAAGCAACAGGGTGCCTTCGG 898  
QY 3659 GACAGCG-CCATCAATACCGC-GGCTTCATCACCAAGG---CAGCGCAGCTGACGTCC 3713  
Db 899 GACAGCGGCTATCATATCCCGGGTTTCATCATCAACGAGGAGCCAGCTGACGTCC 958  
QY 3714 TGTACAAAG 3722  
Db 959 TGACAAAG 967

RESULT 14  
BM910785 1022 bp mRNA linear EST 12-MAR-2002  
LOCUS AGENCOURT\_6617197 NIH\_MGC\_98 Homo sapiens cDNA clone IMAGE:5454337  
DEFINITION 5', mRNA sequence.  
ACCESSION BM910785  
VERSION BM910785.1 GI:19361164  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE 1 (bases 1 to 1022)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: L1CM1949 row: f column: 02  
 High quality sequence start: 12  
 High quality sequence stop: 554.  
 Location/Qualifiers

## FEATURES

source

1..1022  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:545437"  
 /issue\_type="astrocytoma grade IV, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_11b="NIH\_MGC\_98"  
 /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Library constructed by Ling Hong  
 in the Laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 9.2%; Score 787.6; DB 3; Length 1022;  
 Best Local Similarity 91.2%; Pred. No. 7.1e-137;  
 Matches 870; Conservative 0; Mismatches 79; Indels 5; Gaps 3;

6452 CACCCACAGCAGCTCAGCGCAGCCCGCCGCTTACCTCCCTGCGGCGCAGC 6511  
 13 CACCCACAGCAGCTCAGCGCAGCCCGCCGCTTACCTCCCTGCGGCGCAGC 72  
 6512 TCCCCGCTCTGAGACTCGCGCGCCACCCAGTGAACCTTACCTCCGCGCCGAGCAT 6571  
 73 TCCCCGCTCTGAGACTCGCGCGCCACCCAGTGAACCTTACCTCCGCGCCGAGCAT 132  
 6572 GGTGCCCCGCGCGCTGCGCTCCCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 6631  
 133 GGTGCCCCGCGCGCTGCGCTCCCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 192  
 6632 AAGAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 6691  
 193 AAGAGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 252  
 6692 ATGACGAGCAGGAGCACTCCGCGAGTGTGTATCCCGCTGTCGATCCGGATGGGAA 6751  
 253 ATGACGAGCAGGAGCACTCCGCGAGTGTGTATCCCGCTGTCGATCCGGATGGGAA 312  
 6752 CAGACGAGCAGGAGGAGTGGGCTTCAGAGTCTCCAGGAGCAACCAAGCCAGCC 6811  
 313 CAGACGAGCAGGAGGAGTGGGCTTCAGAGTCTCCAGGAGCAACCAAGCCAGCC 372  
 6812 TTCTTACCAAGCTACCGAGAGCAATCCCGCATGTGTATCCCAAGAGCAAGATC 6871  
 373 TTCTTACCAAGCTACCGAGAGCAATCCCGCATGTGTATCCCAAGAGCAAGATC 432  
 6872 AACAGAAAGCTGAACCAACCAACCGAGATGAGCTGAATATCAATATGACGAGCTGG 6931  
 433 AACAGAAAGCTGAACCAACCAACCGAGATGAGCTGAATATCAATATGACGAGCTGG 492  
 6932 ACGGAGATCTTCAATATGCTCCGCATCACCGAGAACAGGCTTATGACCTATGAAAGCC 6991  
 493 ACGGAGATCTTCAATATGCTCCGCATCACCGAGAACAGGCTTATGACCTATGAAAGCC 552

QY 6992 GCGGTGACAGAAATGCGCAGACCAACATGAGGCTGAGGCGCATATTAAGAAAGCACTC 7051  
 DB 553 GCGGTGACAGAAATGCGCAGACCAACATGAGGCTGAGGCGCATATTAAGAAAGCACTC 612  
 QY 7052 ATGGGTAATATGACCAAGTGGGAAGATCCCGCGCTCAGCGCCCAATGCTTTAAACCT 7111  
 DB 613 ATGGGTAATATGACCAAGTGGGAAGATCCCGCGCTCAGCGCCCAATGCTTTAAACCT 672  
 QY 7112 CTGAATGCGAGTGCAGGCTGCGCGCTGCTATGCGCATTAACCGCTCTACGAGCAACT 7171  
 DB 673 CTGAATGCGAGTGCAGGCTGCGCGCTGCTATGCGCATTAACCGCTCTACGAGCAACT 732  
 QY 7172 GACACACACTACCTCGCAGAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7231  
 DB 733 GACACACACTACCTCGCAGAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 792  
 QY 7232 AGCCGAAAGCCAG-TCGCCGCGCCCGGAGCTGAGCATCTGAGGA-CCGAGCACCTTC 7287  
 DB 793 AACCTTAAAGCCAGTCCCGCGCGCTGAGCATCTGAGGAACCGGCAACCTTC 852  
 QY 7288 TGTCTCTCAGTGCCT-CGAGGAGAGTCAACCGCGGAGCGCTCACCACCGCG 7346  
 DB 853 TGTCTCTCAGGAGGCGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 912  
 QY 7347 TGTGAGAGGAGGAGGCGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7400  
 DB 913 CCGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 966

RESULT 15  
 BQ711119 875 bp mRNA linear EST 16-JUL-2002  
 LOCUS BQ711119  
 DEFINITION AGENCOURT\_7977236 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:6215773  
 ACCESSION BQ711119  
 VERSION BQ711119.1 GI:21850018  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.  
 1 (bases 1 to 875)  
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: Dr. Mark Watson  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1CM2385 row: d column: 14  
 High quality sequence stop: 668.  
 Location/Qualifiers

## FEATURES

source

1..875  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6215773"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_11b="NIH\_MGC\_113"  
 /note="Organ: spleen; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Library constructed by Ling Hong in the  
 Laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a

## ORIGIN

NIH\_MGC Library."

Query Match 9.1%; Score 782; DB 5; Length 875;

Best Local Similarity 93.4%; Pred No. 7.9e-136; Matches 808; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

QY	6795	CCAGCCAGCCCGCAGCCTTCTTCAAGCAAGCTGACCGAGGCAACTCCGCATGTCAAGT	6854
DB	1	CCAGCCAGCCCGCAGCCTTCTTCAAGCAAGCTGACCGAGGCAACTCCGCATGTCAAGT	60
QY	6855	CCAAGAACGAGATCAACAAGAACTGAAACCCCAACCGGAATGAGCTGAATACA	6914
DB	61	CCAAGAACGAGATCAACAAGAACTGAAACCCCAACCGGAATGAGCTGAATACA	120
QY	6915	ATATCAGCAGCCTGGAGCGAGATCTTCAATATGCCCCCATGACCGGAAACAGGCTTAA	6974
DB	121	ATATCAGCAGCCTGGAGCGAGATCTTCAATATGCCCCCATGACCGGAAACAGGCTTAA	180
QY	6975	TGACCTATAGAAAGCAGGCGGTGACAGAAACATGCCAGCAACCAATGAGGCTGAGGCCA	7034
DB	181	TGACCTATAGAAAGCAGGCGGTGACAGAAACATGCCAGCAACCAATGAGGCTGAGGCCA	240
QY	7035	TAATTAGAAAGGCACTCATGGTAAATATGACCAATGGGAAAGTCCCCGCGCTCAGCG	7094
DB	241	TAATTAGAAAGGCACTCATGGTAAATATGACCAATGGGAAAGTCCCCGCGCTCAGCG	300
QY	7095	CCAATGCTTTTAAACCTCTGAATGCGAGCTGCGCTGCTATGCCCCATTAACCG	7154
DB	301	CCAATGCTTTTAAACCTCTGAATGCGAGCTGCGCTGCTATGCCCCATTAACCG	360
QY	7155	CTGCTGACGAGCGAGTGAACCAACACTCACTGCGCCAGTGGCGGCGGAAAGCCAAAG	7214
DB	361	CTGCTGACGAGCGAGTGAACCAACACTCACTGCGCCAGTGGCGGCGGAAAGCCAAAG	420
QY	7215	TCTCTGCGAGAACCCGAGAGCCGAAAGTCCCCGGCCCCGGGCTTGGCATCTGGGG	7274
DB	421	TCTCTGCGAGAACCCGAGAGCCGAAAGTCCCCGGCCCCGGGCTTGGCATCTGGGG	480
QY	7275	ACCGGCGAACCTCTGTCTCTCTGAGCACTCGGAGGAGACTGCAACCGCCGGAACGCCGC	7334
DB	481	ACCGGCGAACCTCTGTCTCTCTGAGCACTCGGAGGAGACTGCAACCGCCGGAACGCCGC	540
QY	7335	TCACCAACCGCGTGTGGAGAGACAGGCCCTCGTCCGCAAGTTTCAAGCCATTCCCTTACA	7394
DB	541	TCACCAACCGCGTGTGGAGAGACAGGCCCTCGTCCGCAAGTTTCAAGCCATTCCCTTACA	600
QY	7395	ACCCCTGATCATGCGGCTGACAGGCGGGTGTATGGCTTCCCAACCCCAACCGGGCTTCC	7454
DB	601	ACCCCTGATCATGCGGCTGACAGGCGGGTGTATGGCTTCCCAACCCCAACCGGGCTTCC	660
QY	7455	CCGCGGGGAGCGGGGCCCTCGCTGAGCCCCACCAAGCTTGGGACGAGAGCCCAAGCCAC	7514
DB	661	CCGCGGGGAGCGGGGCCCTCGCTGAGCCCCACCAAGCTTGGGACGAGAGCCCAAGCCAC	720
QY	7515	TGCTCTGCTCGAGTACAGACACTCTCCGACAGCGAGTGACTCAGAAACAGGGCGGGGGG	7574
DB	721	TGCTCTGCTCGAGTACAGACACTCTCCGACAGCGAGTGACTCA-AAAAGGCGNCGGGG	779
QY	7575	GGGCGGGCGGTGEBUSPERLUCTCAGTCCAGCGAGCCACAGAAACGGCCCTGCGAGAG	7634
DB	780	GGGCGGGGNNCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	839
QY	7635	CGGGGCGGCTGCGCACTCCCCCAAC	7659
DB	840	GGCGNCGCGCGCGCGCCCCCNC	864

Search completed: March 12, 2006, 19:33:44  
Job time : 19983 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2006, 04:54:20 ; Search time 3019 Seconds  
(without alignments)  
18921.193 Million cell updates/sec

Title: US-09-522-753-4  
Perfect score: 8564.8  
Sequence: 1 catgctgggctccacacagc.....caaaaaaaaaaaaaaaaa 8571

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_21.\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001s:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003s:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2004s:\*  
12: geneseqn2004bs:\*  
13: geneseqn2004cs:\*  
14: geneseqn2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8541	99.7	8561	12	ADG86298 Human SMR
2	8541	99.7	8561	12	ADN04304 Antipsoi
3	8495.2	99.2	8564	3	AACT4783 Human ORP
4	8459.2	98.8	8667	11	ACN44283 Human mRN
5	8423.4	98.3	8533	13	ADG84524 Human tum
6	8423.4	98.3	8533	13	ACN39603 Tumour-ab
7	8324.6	97.2	8686	8	ACA62249 enco
8	8324.6	97.2	8686	10	ADL13811 Osteoarth
9	8324.6	97.2	8686	12	ADG86290 Human SMR
10	8324.6	97.2	8686	12	ADG18820 Human sof
11	8226.8	96.1	9053	12	ADL12577 Human ble
12	8226.8	96.1	9079	12	ADQ23294 Human sof
13	7554	75.4	7554	12	ADJ92815 Human co-
14	7337.6	85.7	7524	10	ADL13812 Osteoarth
15	7334.6	85.6	7521	8	ACA62250 Human nuc
16	5455	63.7	5989	6	ABK84305 Human CDN
17	5037.2	58.8	8544	8	ACA62451 enco
18	5037.2	58.8	8544	14	ADG1802 Murine NC
19	4797	56.0	7386	8	ACA62452 Mouse nuc

20	4543	53.0	7534	11	ACN44281	ACN44281 Mouse mRN
21	2618.6	30.6	2930	10	ADC35130	ADC35130 Human bre
22	2618.6	30.6	2930	12	ADG86301	ADG86301 Human SMR
23	1138.6	13.3	23380	11	ACN44282	ACN44282 Human gen
24	1132.4	13.2	220756	12	ADG86300	ADG86300 Human SMR
25	949.4	11.1	1027	3	AACT8781	AACT8781 Human pan
26	898.4	10.5	956	3	AAA98862	AAA98862 Human pro
27	898.4	10.5	956	3	AAA98144	AAA98144 Human pro
28	898.4	10.5	956	6	ADG29973	ADG29973 Human pha
29	760.6	8.9	2336	10	ADG31306	ADG31306 Human dia
30	652.8	7.6	752	10	ADL62546	ADL62546 Human apo
31	650	7.6	650	8	ACA57401	ACA57401 Human adl
32	630	7.4	7914	12	ADG83959	ADG83959 Human tum
33	630	7.4	7914	13	ADG87683	ADG87683 Human tum
34	630	7.4	7940	8	ABZ34833	ABZ34833 Coding se
35	630	7.4	7940	13	ADG89791	ADG89791 Antagonis
36	630	7.4	7940	14	ADZ49335	ADZ49335 Insulin B
37	629.6	7.4	7780	3	AAA60630	AAA60630 HNRKR nuc
38	626.8	7.3	2745	12	ADJ92816	ADJ92816 Human co-
39	626.8	7.3	2745	12	ADJ92816	ADJ92816 Human HNR
40	555	6.5	555	8	ACA57524	ACA57524 Human adl
41	542.4	6.3	718	10	ADG76358	ADG76358 Human BSK
42	534	6.2	534	10	ADG79324	ADG79324 Leukaemia
43	494.6	5.8	2914	13	ADG07585	ADG07585 Full leng
44	458.8	5.4	527	13	ADU13876	ADU13876 Solid tum
45	432.8	5.1	91141	11	ACN44280	ACN44280 Mouse gen

## ALIGNMENTS

RESULT 1	ADG86298	ADG86298 standard; cDNA; 8561 BP.
ID	ADG86298	
AC	ADG86298	
XX		
DT	11-MAR-2004	(first entry)
XX		
DE	Human SMRT encoding cDNA SEQ ID NO:12.	
XX		
KW	SMRT, silencing mediator for retinoid and thyroid hormone action;	
KW	SMRT inhibitor; cytosolic; antiinflammatory; antiarthritic;	
KW	antirheumatic; antisense therapy; inflammatory disorder;	
KW	rheumatoid arthritis; hyperproliferative disorder; cancer; leukaemia;	
KW	breast cancer; human; gene; 88.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	2..7555
FT		/*tag= a
FT		/product= "SMRT"
XX		
PN	WO2003106645-A2.	
XX		
PD	24-DEC-2003.	
XX		
PF	17-JUN-2003; 2003WO-US018923.	
XX		
PR	17-JUN-2002; 2002US-00174014.	
XX		
PA	(ISIS-) ISIS PHARM INC.	
XX		
PI	Bennett CF, Freier SM, Doble KW;	
XX		
DR	WPI; 2004-082184/08.	
XX		
DR	P-PSDB; ADG86299.	
XX		
DR	GENBANK; NM_006312.	
XX		
PT	Novel antisense compound targeted to nucleic acid encoding SMRT	
XX	(silencing mediator for retinoid and thyroid hormone action), useful for	
PT	treating animal having disease associated with SMRT such as cancer,	

PT rheumatoid arthritis.  
XX  
PS Example 15; SEQ ID NO 12; 260pp; English.  
XX  
CC The present invention describes a compound (I) 8-50 nucleobases in length  
CC targeted to a nucleic acid molecule encoding SMRT (silencing mediator for  
CC retinoid and thyroid hormone action), where (I) specifically hybridises  
CC with the nucleic acid molecule encoding SMRT and inhibits expression of  
CC SMRT. (I) specifically hybridises with at least 8-nucleobase portion of a  
CC preferred target region on nucleic acid molecule encoding SMRT. Also  
CC described is a composition (II) comprising (I) and a carrier or diluent.  
CC (I) and (II) have cytostatic, antiinflammatory, antiarthritic and  
CC antirheumatic activities, and can be used in antisense therapy, and as  
CC SMRT expression inhibitors. (II) is useful for inhibiting the expression  
CC of SMRT in cells or tissues. (I) is also useful for treating an animal  
CC having a disease or condition associated with SMRT, e.g., inflammatory  
CC disorder such as rheumatoid arthritis, or a hyperproliferative disorder  
CC such as cancer chosen from leukaemia and breast cancer, by inhibiting the  
CC expression of SMRT. (I) is useful for diagnostics, therapeutics,  
CC prophylaxis and as research reagents and kits. The present sequence  
CC encodes human SMRT, which is used in an example from the present  
CC invention. N.B. The present sequence is designated as SEQ ID NO:11 in  
CC example 15 but corresponds to SEQ ID NO:12 in the sequence listing.  
XX  
SQ Sequence 8561 BP, 1862 A, 3033 C, 2525 G, 1141 T, 0 U, 0 Other;  
Query Match 99.7%; Score 8541; DB 12; Length 8561;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 8560; Conservative 1; Mismatches 0; Indels 10; Gaps 1;  
QY 1 CATGTGGGCTCCACACAGCTTTGTGGACAGACGTGAGAGGCGCACTGAGCCCCGCTACCC 60  
DB 1 CATGTGGGCTCCACACAGCTTTGTGGACAGACGTGAGAGGCGCACTGAGCCCCGCTACCC 60  
QY 61 GCCCCACAGCCCTTCTCTACCCGAGTGCAGATTCGCCCGGAGGCAACAGAGAGTGGGCTCT 120  
DB 61 GCCCCACAGCCCTTCTCTACCCGAGTGCAGATTCGCCCGGAGGCAACAGAGAGTGGGCTCT 120  
QY 121 GGAGTACACGACACACTCCCGGACTATGCTCCCACTGTGCGCGGGCTCATATCCA 180  
DB 121 GGAGTACACGACACACTCCCGGACTATGCTCCCACTGTGCGCGGGCTCATATCCA 180  
QY 181 GCCCAGCGGCGAGAGGCTCTCCCTGTCTGTAGTTCCAGCCCGGAGATGAACGGTCCCA 240  
DB 181 GCCCAGCGGCGAGAGGCTCTCCCTGTCTGTAGTTCCAGCCCGGAGATGAACGGTCCCA 240  
QY 241 GGAGTCCACCTGCGCGGACGATCCCACTCATATCTGCGCGGAGTGGGAAATGCAGAGAT 300  
DB 241 GGAGTCCACCTGCGCGGACGATCCCACTCATATCTGCGCGGAGTGGGAAATGCAGAGAT 300  
QY 301 GGAGTTCAATTGAAACAAAGCGCCTCGGCTAGAGCTGCTGACCCCTGCTGCGACC 360  
DB 301 GGAGTTCAATTGAAACAAAGCGCCTCGGCTAGAGCTGCTGACCCCTGCTGCGACC 360  
QY 361 GTCACTCCCTGTGCGCAACGAGCAAGCTTGCGGATCTGAAGACCTCAACCAAGACCGTAG 420  
DB 361 GTCACTCCCTGTGCGCAACGAGCAAGCTTGCGGATCTGAAGACCTCAACCAAGACCGTAG 420  
QY 421 CCTGACGGGCAAGCTGGAACCGGTGTCTCCCGGACCGCCCGGACACTGACCTGAGGT 480  
DB 421 CCTGACGGGCAAGCTGGAACCGGTGTCTCCCGGACCGCCCGGACACTGACCTGAGGT 480  
QY 481 GGAGCTGTGCGCGCACGGCTGTCCAAAGGAGAGCTGATCCAGAACATGGAACCGGTGA 540  
DB 481 GGAGCTGTGCGCGCACGGCTGTCCAAAGGAGAGCTGATCCAGAACATGGAACCGGTGA 540  
QY 541 CCGAGAGATCACCATGTGTAGACAGAGATCTTAAGCTGAAGAGACACAAACAGCT 600  
DB 541 CCGAGAGATCACCATGTGTAGACAGAGATCTTAAGCTGAAGAGACACAAACAGCT 600  
QY 601 GGAGAGAGAGGCTGCGCAAGCGCCCGGAGCTGAGAAAGCCGTGTCAACCGCGCCATGCA 660  
DB 601 GGAGAGAGAGGCTGCGCAAGCGCCCGGAGCTGAGAAAGCCGTGTCAACCGCGCCATGCA 660

QY 661 GTCGAAGACCGCAGCCTGTGTGACAGATCATACGACGAGAACCGGAAAGGCTGAAGC 720  
DB 661 GTCGAAGACCGCAGCCTGTGTGACAGATCATACGACGAGAACCGGAAAGGCTGAAGC 720  
QY 721 TGCACATGGAATTTGTGAAGGCTTGGGCGCCAGGTGAGCTGCGCTGTACACACAGCC 780  
DB 721 TGCACATGGAATTTGTGAAGGCTTGGGCGCCAGGTGAGCTGCGCTGTACACACAGCC 780  
QY 781 CTCGACACCCGGGCGATTCATGAGAACATCAAAATTAACGAGGATGCGGAAAGGT 840  
DB 781 CTCGACACCCGGGCGATTCATGAGAACATCAAAATTAACGAGGATGCGGAAAGGT 840  
QY 841 AATCTTGATCTTCAAGAGAGAGATCAAGCTCGGAAACATGGAAGACAGAAATTTGCGCA 900  
DB 841 AATCTTGATCTTCAAGAGAGAGATCAAGCTCGGAAACATGGAAGACAGAAATTTGCGCA 900  
QY 901 GCGCTATGACCAAGCTCATGAGAGCCTTG3AAAAAAGGTGAGCGCATCGAAAACACCC 960  
DB 901 GCGCTATGACCAAGCTCATGAGAGCCTTG3AAAAAAGGTGAGCGCATCGAAAACACCC 960  
QY 961 GCGCGGCGGCGCAAGAGAGAGAGTGCAGAGTACTACGAAAGCAAGTTCCCTGAGAT 1020  
DB 961 GCGCGGCGGCGCAAGAGAGAGAGTGCAGAGTACTACGAAAGCAAGTTCCCTGAGAT 1020  
QY 1021 CCGCAAGACGCGCAGCTGCAAGAGCGCATGCAAGACAGGCTGGGCAACGCGGGCAGTTGG 1080  
DB 1021 CCGCAAGACGCGCAGCTGCAAGAGCGCATGCAAGACAGGCTGGGCAACGCGGGCAGTTGG 1080  
QY 1081 GCTGTTCATGTCGCGCGCCCGCAGACAGAGGTTCAGAGATCATCATGATGCGCTCTC 1140  
DB 1081 GCTGTTCATGTCGCGCGCCCGCAGACAGAGGTTCAGAGATCATCATGATGCGCTCTC 1140  
QY 1141 AGAGCAGAGAAACCTGAGAAAGCAGATGCGCAGCTGCGCGGATCCGCGCCATCTGTA 1200  
DB 1141 AGAGCAGAGAAACCTGAGAAAGCAGATGCGCAGCTGCGCGGATCCGCGCCATCTGTA 1200  
QY 1201 CGACGCTGACACGACCGCATCAAGTTTCATCAACATGAACGGGCTTATGGCGACCCAT 1260  
DB 1201 CGACGCTGACACGACCGCATCAAGTTTCATCAACATGAACGGGCTTATGGCGACCCAT 1260  
QY 1261 GAAAGTGTACAAAGACCGCAGGTCAATGAATGTGAGTGAACGAGAGAGAACTTT 1320  
DB 1261 GAAAGTGTACAAAGACCGCAGGTCAATGAATGTGAGTGAACGAGAGAGAACTTT 1320  
QY 1321 CCGGAGAAAGTTCAATGACGACATCCCAAGACTTTGGCTGATCGCATATTCCTGGAAG 1380  
DB 1321 CCGGAGAAAGTTCAATGACGACATCCCAAGACTTTGGCTGATCGCATATTCCTGGAAG 1380  
QY 1381 GAAAGCAAGTGGCTGAGTGCCTCTCTATTACTACTCTGACTTAAGAAATGAGAACTATTA 1440  
DB 1381 GAAAGCAAGTGGCTGAGTGCCTCTCTATTACTACTCTGACTTAAGAAATGAGAACTATTA 1440  
QY 1441 GAGCTGTGTAGACGAGACTATTCGCGCGCGGCAAGACCGCAGCAACACAGCAGCA 1500  
DB 1441 GAGCTGTGTGTAGACGAGACTATTCGCGCGCGGCAAGACCGCAGCAACACAGCAGCA 1500  
QY 1501 GCAGCAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1560  
DB 1501 GCAGCAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1560  
QY 1561 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
DB 1561 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
QY 1621 CGACAAAGAGACCTCTCTCAAGAGAGAGACAGACACCTTCAGGGGAGAGACAGACGA 1680  
DB 1621 CGACAAAGAGACCTCTCTCAAGAGAGAGACAGACACCTTCAGGGGAGAGACAGACGA 1680  
QY 1681 GAAAGAGGCTGTGCGCTCCAAAGCGCGCAAAACTGSCCAACGCGGAGAGAGCGCAAGG 1740  
DB 1681 GAAAGAGGCTGTGCGCTCCAAAGCGCGCAAAACTGSCCAACGCGGAGAGAGCGCAAGG 1740



QY 1741 CCGCATACCCGCTCATGTGCTAATGAGCCAAACGGAGAGGCCATCAACCCCGCAGCA 1800  
DB 1741 CCGCATACCCGCTCATGTGCTAATGAGCCAAACGGAGAGGCCATCAACCCCGCAGCA 1800  
QY 1801 GAGCGCGAGCTGGCTCCATGGAAGTGAATGAGAGTTCTCGCTGGAAGAGAGAAAT 1860  
DB 1801 GAGCGCGAGCTGGCTCCATGGAAGTGAATGAGAGTTCTCGCTGGAAGAGAGAAAT 1860  
QY 1861 GGAAGCAGCCAAAGAGTCTCTGGAACAAGCGCGCACTGGTCGGCCGATCGCCGGAT 1920  
DB 1861 GGAAGCAGCCAAAGAGTCTCTGGAACAAGCGCGCACTGGTCGGCCGATCGCCGGAT 1920  
QY 1921 GGTGGGCTCCAGACTGTGTGCGAGTGTAAAGACTTCTCACTTCACTCAAGAGAGCA 1980  
DB 1921 GGTGGGCTCCAGACTGTGTGCGAGTGTAAAGACTTCTCACTTCACTCAAGAGAGCA 1980  
QY 1981 GAACTCTGATGAGATCTTGACAGCACAAGCTGAAGATGAGAGAGAGAGAGAGCGCG 2040  
DB 1981 GAACTCTGATGAGATCTTGACAGCACAAGCTGAAGATGAGAGAGAGAGAGAGCGCG 2040  
QY 2041 GAGGAAGAAAGAAAGCGCGCGCGCGCGCAGCGAGGAGGCTGCAATCCCGCGGTGT 2100  
DB 2041 GAGGAAGAAAGAAAGCGCGCGCGCGCGCAGCGAGGAGGCTGCAATCCCGCGGTGT 2100  
QY 2101 GAGGATGAGAGATGAGAGCGCTCGGCGCTGAGCGGAAATGAGAGAGAGATGCTGAGGA 2160  
DB 2101 GAGGATGAGAGATGAGAGCGCTCGGCGCTGAGCGGAAATGAGAGAGAGATGCTGAGGA 2160  
QY 2161 GGTGAAGCTTATCATGCTCTGGGAATGAGGTGCCAAGGGGAAATGCAGTGGCCAGC 2220  
DB 2161 GGTGAAGCTTATCATGCTCTGGGAATGAGGTGCCAAGGGGAAATGCAGTGGCCAGC 2220  
QY 2221 CACTGTCACAAACAGTCTGACACACGAGAGATCCCTCTCCCTCAACTGAGGCGCCCA 2280  
DB 2221 CACTGTCACAAACAGTCTGACACACGAGAGATCCCTCTCCCTCAACTGAGGCGCCCA 2280  
QY 2281 GGAACAAGGGCAGAAATGAGGCCCAAGCGCCAGCCACTTGAGCGCGAGCAGCCCGC 2340  
DB 2281 GGAACAAGGGCAGAAATGAGGCCCAAGCGCCAGCCACTTGAGCGCGAGCAGCCCGC 2340  
QY 2341 AAGGCCACCCACCCCAACGAGAGACATCCCGGGCCCCCATTTGAGCCCAACCCGCGCTC 2400  
DB 2341 AAGGCCACCCACCCCAACGAGAGACATCCCGGGCCCCCATTTGAGCCCAACCCGCGCTC 2400  
QY 2401 TGAAGGCAACGAGCGCCCTAACGCGCCCAACAGACCCCAATGCGCTCTGCACTCTCC 2460  
DB 2401 TGAAGGCAACGAGCGCCCTAACGCGCCCAACAGACCCCAATGCGCTCTGCACTCTCC 2460  
QY 2461 TGTGTCTCCCAAGAGAGAAAGAGAGAGACCGAGCAGCGCCCGCAGTGGAGAGG 2520  
DB 2461 TGTGTCTCCCAAGAGAGAAAGAGAGAGACCGAGCAGCGCCCGCAGTGGAGAGG 2520  
QY 2521 GAGAGAGCAGAAAGCCCCCGCGCTGAGAGCTGAGAGTGAACAAGGAAAGCCGAGGA 2580  
DB 2521 GAGAGAGCAGAAAGCCCCCGCGCTGAGAGCTGAGAGTGAACAAGGAAAGCCGAGGA 2580  
QY 2581 GCGCGTCAGAGAGATGACCGAGAGAACCGAGAGAGGGCGCGCAAGGGCAAGAACCC 2640  
DB 2581 GCGCGTCAGAGAGATGACCGAGAGAACCGAGAGAGGGCGCGCAAGGGCAAGAACCC 2640  
QY 2641 GAGAGCGCTGAGAGCAGCGCGAGGGGCGCTCAAGGCAAGAGAAAGAGAGGGCGGAG 2700  
DB 2641 GAGAGCGCTGAGAGCAGCGCGAGGGGCGCTCAAGGCAAGAGAAAGAGAGGGCGGAG 2700  
QY 2701 CCGCAGGGCACAACCTGCAAGAGCTCGGCGCGCCCCCAGAGACGGAATCCAGTCTAC 2760  
DB 2701 CCGCAGGGCACAACCTGCAAGAGCTCGGCGCGCCCCCAGAGACGGAATCCAGTCTAC 2760  
QY 2761 CCGCAGTGAAGAGAGTGAATGAAGCCGAGAGGGCGCGCAAGAACCGGCTGTGTCGCC 2820  
DB 2761 CCGCAGTGAAGAGAGTGAATGAAGCCGAGAGGGCGCGCAAGAACCGGCTGTGTCGCC 2820  
QY 2821 AAGGCCAAGCTCTCAACCCCGACTGGCAACCCCGGCAATGTGCTCAACCCCAAGAGCC 2880

DB 2821 AAGGCCAAGCTCTCAACCCCGACTGGCAACCCCGGCAATGTGCTCAACCCCAAGAGCC 2880  
QY 2881 ACTGACCTGAAAGCACTGAAAGAGAGAGGGCTGCAATCCCCCATTCAGGTCAACAA 2940  
DB 2881 ACTGACCTGAAAGCACTGAAAGAGAGGGCTGCAATCCCCCATTCAGGTCAACAA 2940  
QY 2941 AGTCATGAGCCCCCGGAGAGACGAGCTCCCAACCAAGCAGTCTCCCAAGCCCAACC 3000  
DB 2941 AGTCATGAGCCCCCGGAGAGACGAGCTCCCAACCAAGCAGTCTCCCAAGCCCAACC 3000  
QY 3001 GGCACGCAAAACCTGACGCGGAGAGAGCGCCCTCAGACACTGCGAGAGCCCGC 3060  
DB 3001 GGCACGCAAAACCTGACGCGGAGAGAGCGCCCTCAGACACTGCGAGAGCCCGC 3060  
QY 3061 GGGCAAGACAGAGAGCGCGGCAACCCCGCGCAAGAGAGGCTTGGCAAGCGAGCGCA 3120  
DB 3061 GGGCAAGACAGAGAGCGCGGCAACCCCGCGCAAGAGAGGCTTGGCAAGCGAGCGCA 3120  
QY 3121 GAACTGCTGAGGAGACCCCTTGTGAGACTTCCGAGCTGACCCTTCCCGTGGCCCGC 3180  
DB 3121 GAACTGCTGAGGAGACCCCTTGTGAGACTTCCGAGCTGACCCTTCCCGTGGCCCGC 3180  
QY 3181 TGAAGTGAATCAAGGCTCCCGGATGCCCGAGACCCCTGACGCTTCTTACGCTCAC 3240  
DB 3181 TGAAGTGAATCAAGGCTCCCGGATGCCCGAGACCCCTGACGCTTCTTACGCTCAC 3240  
QY 3241 TGTGACCCCACTGCCCCCTGGGCTCCATGACACTGCCCCCGGCTCTGCGGCGCCAC 3300  
DB 3241 TGTGACCCCACTGCCCCCTGGGCTCCATGACACTGCCCCCGGCTCTGCGGCGCCAC 3300  
QY 3301 CACCATCTCAACCGCGCTCCCTCATCTCTCTGCAAGACACCCAGCGCTCTCGAGAG 3360  
DB 3301 CACCATCTCAACCGCGCTCCCTCATCTCTCTGCAAGACACCCAGCGCTCTCGAGAG 3360  
QY 3361 GCAAAATAGTGCATCTCCCAAGAAATGTGAGTCAAGTCAAGTCCGTAATCAAGACA 3420  
DB 3361 GCAAAATAGTGCATCTCCCAAGAAATGTGAGTCAAGTCAAGTCCGTAATCAAGACA 3420  
QY 3421 TGCCAAGGCCCCGGTGGGCTCTGTCAACATGAGGCTGCCCCCTGCAATGAGACCCCAAAA 3480  
DB 3421 TGCCAAGGCCCCGGTGGGCTCTGTCAACATGAGGCTGCCCCCTGCAATGAGACCCCAAAA 3480  
QY 3481 GCTGGACCCCTCAAGGAGTGAAGAGAGAGCAGTGTCCCAACGAGGCGCAAGGCTGG 3540  
DB 3481 GCTGGACCCCTCAAGGAGTGAAGAGAGAGCAGTGTCCCAACGAGGCGCAAGGCTGG 3540  
QY 3541 ACCGAGAGCTGGGGGTGCCCAACAGCCAGAGAGGCGTCCGTGAGAGAGGACAAGCTCT 3600  
DB 3541 ACCGAGAGCTGGGGGTGCCCAACAGCCAGAGAGGCGTCCGTGAGAGAGGACAAGCTCT 3600  
QY 3601 GGGCTGAGTTCCGGCGGGAAGCATCAACAAAGCATTTCCAGAGCAGCGGTGCCCCGGA 3660  
DB 3601 GGGCTGAGTTCCGGCGGGAAGCATCAACAAAGCATTTCCAGAGCAGCGGTGCCCCGGA 3660  
QY 3661 CAGCGCATCAATACCGGAGTCCCATCAACCAAGCAGCAGCAGCATGCTCTGTACAA 3720  
DB 3661 CAGCGCATCAATACCGGAGTCCCATCAACCAAGCAGCAGCAGCATGCTCTGTACAA 3720  
QY 3721 GGGCAACATCAACAGATATCGGCGAGAGCAGCCGAGTGTGAGACCGCGCGGAG 3780  
DB 3721 GGGCAACATCAACAGATATCGGCGAGAGCAGCCGAGTGTGAGACCGCGCGGAG 3780  
QY 3781 GGAACGCTTCCCAAGGCGCAGTCAATCAAGAGAGAGAGGCGCAGTCTTGTGCTCA 3840  
DB 3781 GGAACGCTTCCCAAGGCGCAGTCAATCAAGAGAGAGAGGCGCAGTCTTGTGCTCA 3840  
QY 3841 TGAAGGTGCAATGTCTGTGACCAAGGCTCAAGAGAGAGCGGAGAGAGCAGCTCAGAGC 3900  
DB 3841 TGAAGGTGCAATGTCTGTGACCAAGGCTCAAGAGAGAGCGGAGAGAGCAGCTCAGAGC 3900  
QY 3901 CCCCATGAGAGCGCGCGCCCAAGCGCACCTATGACATGATGAGAGCGCGGTGGCGAG 3960

Db 3901 CCCCCATGAGAGCGCGCCGCCCAAGCGCACTATGACATGATGAGAGGCGCGGTGGGAG 3960  
OY AGCCATCTCTCAGCGAGCATGAGAGGTCTCATGGGCGGTGCGCATCCGCGGAGCGACA 4020  
Db 3961 AGCCATCTCTCAGCGAGCATGAGAGGTCTCATGGGCGGTGCGCATCCGCGGAGCGACA 4020  
OY 4021 CAGCCCCACCACTCAAAAGACAGACCAATCCGCGGCTCATCACAAGAGANTCC 4080  
Db 4021 CAGCCCCACCACTCAAAAGACAGACCAATCCGCGGCTCATCACAAGAGANTCC 4080  
OY 4081 TCGGCTCTACGTGAGAGGCAAGAGAGACTAAGTGGGTGGAGAGGCTCTTAAAGG 4140  
Db 4081 TCGGCTCTACGTGAGAGGCAAGAGAGACTAAGTGGGTGGAGAGGCTCTTAAAGG 4140  
OY 4141 GAGAGGCAAGCTCTCGCGCCCGCCCTCACGGGACCTGACCGAGGCTTACAAAGCGCA 4200  
Db 4141 GAGAGGCAAGCTCTCGCGCCCGCCCTCACGGGACCTGACCGAGGCTTACAAAGCGCA 4200  
OY 4201 GGCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGGCTGTGGCCACGGTGAAGGA 4260  
Db 4201 GGCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGGCTGTGGCCACGGTGAAGGA 4260  
OY 4261 GCGCGGCGGCTCTCATTCATGAGATCCCGGCGAGAGGCTGCGGCAACGCGCGAGCTGCC 4320  
Db 4261 GCGCGGCGGCTCTCATTCATGAGATCCCGGCGAGAGGCTGCGGCAACGCGCGAGCTGCC 4320  
OY 4321 CCTGGCCCCGCGCGGCTCAAGAGAGGCTCCATCAGCGAGGCACTCCGCTAAAGTACGA 4380  
Db 4321 CCTGGCCCCGCGCGGCTCAAGAGAGGCTCCATCAGCGAGGCACTCCGCTAAAGTACGA 4380  
OY 4381 CACCGGCGGCTCACACTGGCTCAAAAAGACAGACGATCGCTCCTCATCGGAGGCC 4440  
Db 4381 CACCGGCGGCTCACACTGGCTCAAAAAGACAGACGATCGCTCCTCATCGGAGGCC 4440  
OY 4441 CCGCGGAGCGTTCCCAACCGGTGCAACCGGTGAGATGATGAGCGCGCGGCGCACTGGA 4500  
Db 4441 CCGCGGAGCGTTCCCAACCGGTGCAACCGGTGAGATGATGAGCGCGCGGCGCACTGGA 4500  
OY 4501 AGGTGCTGCTACGAGAGAGGCTGAAGAGCGGCGAGGAGCCGCGAGAGCTCGGGGG 4560  
Db 4501 AGGTGCTGCTACGAGAGAGGCTGAAGAGCGGCGAGGAGCCGCGAGAGCTCGGGGG 4560  
OY 4561 CTTCATTTGCGCGGGCGCCCCGGTCAATTGTGCTTGAAGTGGGTAAAGCGCGGAGAGCCC 4620  
Db 4561 CTTCATTTGCGCGGGCGCCCCGGTCAATTGTGCTTGAAGTGGGTAAAGCGCGGAGAGCCC 4620  
OY 4621 CCTGACCTATAGAGACCAAGGGGCAACCTTTGCGGGCACTCCCAACAGAGTTCGCCGT 4680  
Db 4621 CCTGACCTATAGAGACCAAGGGGCAACCTTTGCGGGCACTCCCAACAGAGTTCGCCGT 4680  
OY 4681 GACCATGCGGAGGAGCCACGCGCGCTGAGAGAGGAGCTTTGTGCTCAGAGAGGACATC 4740  
Db 4681 GACCATGCGGAGGAGCCACGCGCGCTGAGAGAGGAGCTTTGTGCTCAGAGAGGACATC 4740  
OY 4741 CCAAGACCGAAAGCTGACGTGACGCTCTGTGAGATCGGCAAGTCCCGGCAAGACCGT 4800  
Db 4741 CCAAGACCGAAAGCTGACGTGACGCTCTGTGAGATCGGCAAGTCCCGGCAAGACCGT 4800  
OY 4801 GCGCGAGACCAACCCCAACCCCATCTGCGGCTTATAGAGCACTGCTTGGGGGTGATGG 4860  
Db 4801 GCGCGAGACCAACCCCAACCCCATCTGCGGCTTATAGAGCACTGCTTGGGGGTGATGG 4860  
OY 4861 GCGCGAGACCAACCCCAACCCCATCTGCGGCTTATAGAGCACTGCTTGGGGGTGATGG 4860  
Db 4861 GCGCGAGACCAACCCCAACCCCATCTGCGGCTTATAGAGCACTGCTTGGGGGTGATGG 4860  
OY 4921 CATCCCTTGAACGAGCGGCTGCTTAACCTGCCCCGAGCACTGGCCCCAACCCAC 4980  
Db 4921 CATCCCTTGAACGAGCGGCTGCTTAACCTGCCCCGAGCACTGGCCCCAACCCAC 4980  
OY 4981 CTACCCGACCTGTACCAACCTTACTCTATCGCGGCTACCCGAGCAAGGCGGCGCTGGA 5040  
Db 4981 CTACCCGACCTGTACCAACCTTACTCTATCGCGGCTACCCGAGCAAGGCGGCGCTGGA 5040

OY 5041 GAACCGGAGACCATCATCAATGACTATCATGACTTCGAGAGAGATGACCAACACGCGC 5100  
Db 5041 GAACCGGAGAGACCATCATCAATGACTATCATGACTTCGAGAGAGATGACCAACGCGC 5100  
OY 5101 CACCGCATGCGCCAGCGAGCTGATATGCTGAGGGGCTCTTGGCCCCGAGTCTCGCT 5160  
Db 5101 CACCGCATGCGCCAGCGAGCTGATATGCTGAGGGGCTCTTGGCCCCGAGTCTCGCT 5160  
OY 5161 GGCATCTCACTACGCTGGGGGCTCCCGAGGACTATGACCTGTGCTCAAGTGGCAACT 5220  
Db 5161 GGCATCTCACTACGCTGGGGGCTCCCGAGGACTATGACCTGTGCTCAAGTGGCAACT 5220  
OY 5221 GCTGTGCTGTGCTCCCGAGACAGAGACCCAGCGCACCGCATGAGACCGCTTGGCTTA 5280  
Db 5221 GCTGTGCTGTGCTCCCGAGACAGAGACCCAGCGCACCGCATGAGACCGCTTGGCTTA 5280  
OY 5281 CTTCCCAACCGGCGCCAGGCTTTACAGAGCGCGCAAGAGCTGCCACTTCCCACTTCCCAAG 5340  
Db 5281 CTTCCCAACCGGCGCCAGGCTTTACAGAGCGCGCAAGAGCTGCCACTTCCCACTTCCCAAG 5340  
OY 5341 AGGTCAACACACTTGAACAAACCAACCAACGATCGTCTCGAGGGGAGCGAGACCG 5400  
Db 5341 AGGTCAACACACTTGAACAAACCAACCAACGATCGTCTCGAGGGGAGCGAGACCG 5400  
OY 5401 GGAATGAGAGCGGAGACCGGAGATCGGAGATCGGAGGAGGAGGAGGAGGAGGAGGAG 5460  
Db 5401 GGAATGAGAGCGGAGACCGGAGATCGGAGATCGGAGGAGGAGGAGGAGGAGGAGGAG 5460  
OY 5461 GGTGAGAGACCGACCATCTGTGAGAGCTGTGTACAGAGCAAGACGCGGAGAGGCGAG 5520  
Db 5461 GGTGAGAGACCGACCATCTGTGAGAGCTGTGTACAGAGCAAGACGCGGAGAGGCGAG 5520  
OY 5521 CAGCGCGGGGGTGGGGGACAGAGAGCGCGCGGCTCCCATGCTCAATGCGCAACAGA 5580  
Db 5521 CAGCGCGGGGGTGGGGGACAGAGAGCGCGCGGCTCCCATGCTCAATGCGCAACAGA 5580  
OY 5581 CTGCGCCATCTCCCTCGGAGCCAGAGATGCCCTCAGAGAGAGCCAGTGTGCTTACAA 5640  
Db 5581 CTGCGCCATCTCCCTCGGAGCCAGAGATGCCCTCAGAGAGAGCCAGTGTGCTTACAA 5640  
OY 5641 CACAGGCAATGAAGGATCATCAACGCTGTGAGAGCGCGAGACCGCATGCTTGAAGTC 5700  
Db 5641 CACAGGCAATGAAGGATCATCAACGCTGTGAGAGCGCGAGACCGCATGCTTGAAGTC 5700  
OY 5701 CACCTGCACTCTCAACCGGTTGCGCGAGCTGCGCAATTCCTCCACCTGCAACCACTGCC 5760  
Db 5701 CACCTGCACTCTCAACCGGTTGCGCGAGCTGCGCAATTCCTCCACCTGCAACCACTGCC 5760  
OY 5761 ACTGGGCGGCAACCTCGATGGGATCTAACCCTACCTGAGAGCGCGTCTTGCTGCCAA 5820  
Db 5761 ACTGGGCGGCAACCTCGATGGGATCTAACCCTACCTGAGAGCGCGTCTTGCTGCCAA 5820  
OY 5821 GAGGCGCCCCCGGGTGCGCCCGGCGAGAGCGGCGCGAGAGACATCGGCAATGCTTCT 5880  
Db 5821 GAGGCGCCCCCGGGTGCGCCCGGCGAGAGCGGCGCGAGAGACATCGGCAATGCTTCT 5880  
OY 5881 GCGCAAGCCCCAGCGGCTCGGGCTGGAGCGGCGCTCTCCCGCAGAGAGGCTCGGA 5940  
Db 5881 GCGCAAGCCCCAGCGGCTCGGGCTGGAGCGGCGCTCTCCCGCAGAGAGGCTCGGA 5940  
OY 5941 GCGCGGCGCCCTAGTGTGCTCTGCTGTGAGCGACGACCAATGCGCGCAACCCCTGCGAA 6000  
Db 5941 GCGCGGCGCCCTAGTGTGCTCTGCTGTGAGCGACGACCAATGCGCGCAACCCCTGCGAA 6000  
OY 6001 GAACCTGCACTCAACAGCGGAGCGGAGCCGCGGCGGCACTGTGCTCGGCTCGGA 6060  
Db 6001 GAACCTGCACTCAACAGCGGAGCGGAGCCGCGGCGGCACTGTGCTCGGCTCGGA 6060  
OY 6061 CCGGACCGGGAGAAAGACTCAAGATTAACCTTTTCACTCAGAGAACTGGAATCTCGTTTC 6120  
Db 6061 CCGGACCGGGAGAAAGACTCAAGATTAACCTTTTCACTCAGAGAACTGGAATCTCGTTTC 6120

QY 6121 TCTGGGTTACACGGGAGACGTACAGCCGCCGAAAGGGGTGAGCCCTCAGCCCTGTGAG 6180  
DB 6121 TCTGGGTTACACGGGAGACGTACAGCCGCCGAAAGGGGTGAGCCCTCAGCCCTGTGAG 6180  
QY 6181 CTCACCAAGTCTGACCCACGACAAAGGGGCTCCCAAGACCTTGAGAGGTTCCACAAAG 6240  
DB 6181 CTCACCAAGTCTGACCCACGACAAAGGGGCTCCCAAGACCTTGAGAGGTTCCACAAAG 6240  
QY 6241 CCACTGAGAGGGGAGTGGGGCCCAAGGAGCCAGGCCCCGTGAAGCTTGGCCGGGAGGC 6300  
DB 6241 CCACTGAGAGGGGAGTGGGGCCCAAGGAGCCAGGCCCCGTGAAGCTTGGCCGGGAGGC 6300  
QY 6301 CGGCCACTTCCACACCTGCGGCGCTGCTTGAGAGCCAGCCCTGTCAAGCCCGCTGCT 6360  
DB 6301 CGGCCACTTCCACACCTGCGGCGCTGCTTGAGAGCCAGCCCTGTGTCAAGCCCGCTGCT 6360  
QY 6361 CCAAGCCGCCCCAGGGGTCAAAAGGTCAACAGGGGTGTCAACCTGAGCCAGCATCAG 6420  
DB 6361 CCAAGCCGCCCCAGGGGTCAAAAGGTCAACAGGGGTGTCAACCTGAGCCAGCATCAG 6420  
QY 6421 TAGAGTCAACACAGGACTACACCGGGACCAACCAAGAGACTAGCGGACACCCCTGCC 6480  
DB 6421 TAGAGTCAACACAGGACTACACCGGGACCAACCAAGAGACTAGCGGACACCCCTGCC 6480  
QY 6481 CGCCCCCTCTACTCTCTTCCCTGAGGACAGCTGCCGCTCTGAGACTCTCGCGCCACAC 6540  
DB 6481 CGCCCCCTCTACTCTCTTCCCTGAGGACAGCTGCCGCTCTGAGACTCTCGCGCCACAC 6540  
QY 6541 CAGTGACTTCTACCTTCCCGCCCCGAGCAATGATGCCCCGGCCCTGTGACTCCCCCAAG 6600  
DB 6541 CAGTGACTTCTACCTTCCCGCCCCGAGCAATGATGCCCCGGCCCTGTGACTCCCCCAAG 6600  
QY 6601 CCAAGGGGGGCAAGAGTCTTCCAGAGCCCAACAGAGCTGGGTCTTGGGTGGTGTAGGA 6660  
DB 6601 CCAAGGGGGGCAAGAGTCTTCCAGAGCCCAACAGAGCTGGGTCTTGGGTGGTGTAGGA 6660  
QY 6661 CGGATTTGAACTGTGTCCCAACGGAGGGCATGACGGAGCCAGAGCTCCCGAGTGC 6720  
DB 6661 CGGATTTGAACTGTGTCCCAACGGAGGGCATGACGGAGCCAGAGCTCCCGAGTGC 6720  
QY 6721 TGTGTAACCCGCTGTGTACCGGGATGGGGAACAGACGGAGCCAGAGATGGGCTTCAA 6780  
DB 6721 TGTGTAACCCGCTGTGTACCGGGATGGGGAACAGACGGAGCCAGAGATGGGCTTCAA 6780  
QY 6781 GTTCTCAAGGCAACACAGCCAGCCGCAAGCTTCTTCAAGAGCTGACCGAGGCAATC 6840  
DB 6781 GTTCTCAAGGCAACACAGCCAGCCGCAAGCTTCTTCAAGAGCTGACCGAGGCAATC 6840  
QY 6841 CGCATGTCTCAAGTCCAGAGAGGAGATCAACAAGAGCTGAACCCACACCGGAA 6900  
DB 6841 CGCATGTCTCAAGTCCAGAGAGGAGATCAACAAGAGCTGAACCCACACCGGAA 6900  
QY 6901 TGAAGCTGAATACATATCAGCCAGCCTGGGAGGAGATCTTCAATATGCCCATC 6960  
DB 6901 TGAAGCTGAATACATATCAGCCAGCCTGGGAGGAGATCTTCAATATGCCCATC 6960  
QY 6961 CGGAACAGGCTTATGACCTTATGAAGCCAGGCGGTGCAAGAACATGCCAGCAACAT 7020  
DB 6961 CGGAACAGGCTTATGACCTTATGAAGCCAGGCGGTGCAAGAACATGCCAGCAACAT 7020  
QY 7021 GGGGCTGAGAGGCAATATTGAAGGCACTATGGGTAAATATGACAGAGGGAGAGTGC 7080  
DB 7021 GGGGCTGAGAGGCAATATTGAAGGCACTATGGGTAAATATGACAGAGGGAGAGTGC 7080  
QY 7081 CCGCGGCTCAAGGCAATGCTTTTAAACCTCTGAATGCCAGTGCAGCCCTGCGCTGC 7140  
DB 7081 CCGCGGCTCAAGGCAATGCTTTTAAACCTCTGAATGCCAGTGCAGCCCTGCGCTGC 7140  
QY 7141 TATGCCCATTAACGGCTGTGACGGAACGAGTGAACACACTCACTCTGCGCAAGTGGCG 7200  
DB 7141 TATGCCCATTAACGGCTGTGACGGAACGAGTGAACACACTCACTCTGCGCAAGTGGCG 7200  
QY 7201 CGGGAAGCCAAAGTCTGTGGAGACCCAGACCCGAAAGCCAAAGTCCCGGCGCGG 7260

DB 7201 CGGGAAGCCAAAGTCTGTGGAGACCCAGACCCGAAAGCCAAAGTCCCGGCGCGG 7260  
QY 7261 CTTGGCATCTGGGAGACCGGACACCTCTGTCTCTCAATGTCATCTCGAGGGGAATCGCA 7320  
DB 7261 CTTGGCATCTGGGAGACCGGACACCTCTGTCTCTCAATGTCATCTCGAGGGGAATCGCA 7320  
QY 7321 CGCCGGAACGCGCTCAACCAACCGGTGTGGAGAGACAGGCGCTGTCCGAGGTTCCAC 7380  
DB 7321 CGCCGGAACGCGCTCAACCAACCGGTGTGGAGAGACAGGCGCTGTCCGAGGTTCCAC 7380  
QY 7381 GCCATTTCCCTTCAACCCCTGATCATGCGGCTGACAGGCGGTTGATGCTTCCAC 7440  
DB 7381 GCCATTTCCCTTCAACCCCTGATCATGCGGCTGACAGGCGGTTGATGCTTCCAC 7440  
QY 7441 CCAACCGGAGCTTCCCGGGGACGCGGCGCTGTGAGGCGGCGGCGGCGGCGGCGGCGG 7500  
DB 7441 CCAACCGGAGCTTCCCGGGGACGCGGCGCTGTGAGGCGGCGGCGGCGGCGGCGGCGG 7500  
QY 7501 GAGCCCAAGCCACTGTCTGTGCTGAGATACAGACACTTCCGACAGGAGTGAATCAG 7560  
DB 7501 GAGCCCAAGCCACTGTCTGTGCTGAGATACAGACACTTCCGACAGGAGTGAATCAG 7560  
QY 7561 AACAGGCGGGGGGGGGCGGGCGGTGUBERS PERLUGTCAAGTCCAGCGACCAAGAA 7620  
DB 7561 AACAGGCGGGGGGGGGCGGGCGGTGUBERS PERLUGTCAAGTCCAGCGACCAAGAA 7620  
QY 7621 CGGCTCTGACAGAGCGGGGCGGCTGCGGACTCCCCCAACCAAGAAAGAGCCCTGATGC 7680  
DB 7621 CGGCTCTGACAGAGCGGGGCGGCTGCGGACTCCCCCAACCAAGAAAGAGCCCTGATGC 7680  
QY 7681 CGGCTGCGCTGCAATCATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7740  
DB 7681 CGGCTGCGCTGCAATCATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7740  
QY 7741 AACTAAGACTCCCGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7800  
DB 7741 AACTAAGACTCCCGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7800  
QY 7801 GTGCTCGGGAAGGAGAGGAGAGGAGCGGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAG 7860  
DB 7801 GTGCTCGGGAAGGAGAGGAGAGGAGCGGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAG 7860  
QY 7861 ACAAGGCGGCGGAGGCGGCGGAGGAGCCCAAGAGAGATGACAGGACCTCCAGGCAAC 7920  
DB 7861 ACAAGGCGGCGGAGGCGGCGGAGGAGCCCAAGAGAGATGACAGGACCTCCAGGCAAC 7920  
QY 7921 TGCCTCCCGGAATGATTTGGAACCAAGTCTTAACTGAGCTCGAGCCCGCGCTT 7980  
DB 7921 TGCCTCCCGGAATGATTTGGAACCAAGTCTTAACTGAGCTCGAGCCCGCGCTT 7980  
QY 7981 CCTTCGCTCCCATTCGCTTACGCTCTGACAGATGAGCGAGCCCTGTTCAGGCC 8040  
DB 7981 CCTTCGCTCCCATTCGCTTACGCTCTGACAGATGAGCGAGCCCTGTTCAGGCC 8040  
QY 8041 CCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8100  
DB 8041 CCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8100  
QY 8101 GTTCAGGCAAGTGGGCGGCAAAAGGCGAGTGCAGGCTGAGGGGAAACGATGCTCGA 8160  
DB 8101 GTTCAGGCAAGTGGGCGGCAAAAGGCGAGTGCAGGCTGAGGGGAAACGATGCTCGA 8160  
QY 8161 GGAAGTGAAGTGTGTTTTCACACATGTTGCGGACGCGGTGGGAGAAAGGAGATGTA 8220  
DB 8161 GGAAGTGAAGTGTGTTTTCACACATGTTGCGGACGCGGTGGGAGAAAGGAGATGTA 8220  
QY 8221 AATGATGTGTGTTTTCACAGGGATATTTTGTATTAACCTTCAATGAATTAATCAGATGT 8280  
DB 8221 AATGATGTGTGTTTTCACAGGGATATTTTGTATTAACCTTCAATGAATTAATCAGATGT 8280  
QY 8281 TTAAGCAAGAAAGACTTACCAAGTATATGCTGTGCTGCTTGAATCTGCTTACCG 8340

Db	8271	TTACGCAAGGAAGACTTACCCAGATTACTGCTGCTGCTTTTGTACTGCTTACCG	8330
Qy	8341	TTCAAGAGCGCTGTGTCAGAGCGAGTGTGAGTCCCATCATCTGCAAGACCAAGGGGC	8400
Db	8331	TTCAAGAGCGCTGTGTCAGAGCGAGTGTGAGTCCCATCATCTGCAAGACCAAGGGGC	8390
Qy	8401	GGGAGCTGCTGTGTCAGCGCCCGCTGTGTCCTCCCTCCCTCTTGGGCAATGAA	8460
Db	8391	GGGAGCTGCTGTGTCAGCGCCCGCTGTGTCCTCCCTCCCTCTTGGGCAATGAA	8450
Qy	8461	TTTCGATGTGATTTCTGTGCGCCCGCAATTTGGCAAGGTGTGTATTTCTATTTACCA	8520
Db	8451	TTTCGATGTGATTTCTGTGCGCCCGCAATTTGGCAAGGTGTGTATTTCTATTTACCA	8510
Qy	8521	CGTCTTCTAATTTAAAGCGAATTACTCCAAAAAATTTAAAAA	8571
Db	8511	CGTCTTCTAATTTAAAGCGAATTACTCCAAAAAATTTAAAAA	8561
RESULT 2			
ID	ADN04304	standard; cDNA; 8561 BP.	
XX	ADN04304;		
AC	01-JUL-2004	(first entry)	
DT	Antipsoriatic cDNA sequence #352.		
DE	ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.		
XX	Homo sapiens.		
OS	MO2004028479-A2.		
PN	08-APR-2004.		
XX	25-SEP-2003; 2003WO-US030907.		
PF	25-SEP-2002; 2002US-041406P.		
XX	(GENTH ) GENE TECH INC.		
PA	Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;		
XX	Wu TD;		
PI	WPI: 2004-305105/28.		
DR	P-PSDB; ADN04305.		
XX	New PRO nucleic acid or polypeptide, useful for preparing a		
PT	pharmaceutical composition for diagnosing or treating psoriasis in a		
XX	mammal.		
PS	Claim 1; SEQ ID NO 698; 3063bp; English.		
XX	The invention relates to novel polynucleotide and polypeptides for		
CC	treating psoriasis or a sequence having at least 80% identity to the		
CC	above sequences. The nucleic acid is useful for preparing a composition		
CC	for diagnosing or treating psoriasis in a mammal. This sequence		
CC	corresponds to one of the polynucleotides of the invention.		
XX	Sequence 8561 BP; 1862 A; 3033 C; 2525 G; 1141 T; 0 U; 0 Other;		
SQ			
Query Match 99.7%; Score 8541; DB 12; Length 8561;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 8560; Conservative 1; Mismatches 0; Indels 10; Gaps 1;			
Qy	1	CATGTGCGGCTTCACACAGCTTGTGCAAGAGTGTGAGGCGCACTGAGCCCGCTACCC	60
Db	1	CATGTGCGGCTTCACACAGCTTGTGCAAGAGTGTGAGGCGCACTGAGCCCGCTACCC	60
Qy	61	GGCCCAAGCCCTTTCTTACCCAGTGCATGCGCCGAGACGACAGGACGTGGGCTCTT	120

Db	61	GGCCCAAGCCCTTTCTTACCCAGTGCATGCGCCGAGACGACAGGACGTGGGCTCTT	120
Qy	121	GAAGTACAGGACCACTCCCGGACTATAGCTCCACCTGTGCGCGGCTCCATCATCA	180
Db	121	GAAGTACAGGACCACTCCCGGACTATAGCTCCACCTGTGCGCGGCTCCATCATCA	180
Qy	181	GGCCCAAGCGGAGGCGCTTCCCTGTGTCTGAGTTCCAGCCCGGAAATGAACGTTCCA	240
Db	181	GGCCCAAGCGGAGGCGCTTCCCTGTGTCTGAGTTCCAGCCCGGAAATGAACGTTCCA	240
Qy	241	GAAGTCCACCTGCGGCGCAAGTCCACTCATPACCTGCGCGAGCTGGGGAAGTCAAGAT	300
Db	241	GAAGTCCACCTGCGGCGCAAGTCCACTCATPACCTGCGCGAGCTGGGGAAGTCAAGAT	300
Qy	301	GAAGTTCAATTGAAGCAAGCGCCCTCGCTAGAGCTGTGCTGACCCCTGCTGCGACC	360
Db	301	GAAGTTCAATTGAAGCAAGCGCCCTCGCTAGAGCTGTGCTGACCCCTGCTGCGACC	360
Qy	361	GTCAACCCCTGTGCTGCTGACCGGCGACGCTGCGGAACTGAAGACTTACCAAGACCTAG	420
Db	361	GTCAACCCCTGTGCTGCTGACCGGCGACGCTGCGGAACTGAAGACTTACCAAGACCTAG	420
Qy	421	CTGACCGGCAAGCTGGAACCGGTGTCTCCCGCAGCCCGCGCACATGACCCCTGAGCT	480
Db	421	CTGACCGGCAAGCTGGAACCGGTGTCTCCCGCAGCCCGCGCACATGACCCCTGAGCT	480
Qy	481	GGAGCTGTGCTGCGCACCGCTGTCCAGAGAGCTGATCCAGAACATGACCGCGTGA	540
Db	481	GGAGCTGTGCTGCGCACCGCTGTCCAGAGAGCTGATCCAGAACATGACCGCGTGA	540
Qy	541	CCGAGAGATCACCATGTGTAGAGCAGCAGATCTTAACTGAAAGAAAGCAGCAACGCT	600
Db	541	CCGAGAGATCACCATGTGTAGAGCAGCAGATCTTAACTGAAAGAAAGCAGCAACGCT	600
Qy	601	GGAGGAGAGGCTGCTCAAGCCCGCGAGCTGAGAGGCGCGGTCAACCGCGCCATGA	660
Db	601	GGAGGAGAGGCTGCTCAAGCCCGCGAGCTGAGAGGCGCGGTCAACCGCGCCATGA	660
Qy	661	GTGAGAGCAGCGCAGCTGTGTGACAGATCATCAAGCAGAACCCGAGAAAGGCTGAAGC	720
Db	661	GTGAGAGCAGCGCAGCTGTGTGACAGATCATCAAGCAGAACCCGAGAAAGGCTGAAGC	720
Qy	721	TGCATGTGATTTCTGAAAGGCTTGGGCGCCAGGTGAAGCTGCGCTGTACACAGCC	780
Db	721	TGCATGTGATTTCTGAAAGGCTTGGGCGCCAGGTGAAGCTGCGCTGTACACAGCC	780
Qy	781	CTCCGACACCCGCGAGTATCATGAGAACATCAAAATTAACAGGCGAGTGGGAAAGCT	840
Db	781	CTCCGACACCCGCGAGTATCATGAGAACATCAAAATTAACAGGCGAGTGGGAAAGCT	840
Qy	841	AATCTTGTACTTCAAGAGAGGAAATCAGCTTCGAAACATAGAAAGCAAGATTCTGCA	900
Db	841	AATCTTGTACTTCAAGAGAGGAAATCAGCTTCGAAACATAGAAAGCAAGATTCTGCA	900
Qy	901	GGCGTATGACCAAGCTCATGAGGCTTGGAAAAAAGGTGAAGCGCATGAAACACACC	960
Db	901	GGCGTATGACCAAGCTCATGAGGCTTGGAAAAAAGGTGAAGCGCATGAAACACACC	960
Qy	961	GGCGCGGCGGCGCAAGGAGAGCAAGGTGCGGAGTACTACGAAAAAGCATTCCTGAGAT	1020
Db	961	GGCGCGGCGGCGCAAGGAGAGCAAGGTGCGGAGTACTACGAAAAAGCATTCCTGAGAT	1020
Qy	1021	CCGCAAGCAGCGGAGCTGTGAGAGCGCATGACAGAGGCTGGGCGACGCGGCGAGTGG	1080
Db	1021	CCGCAAGCAGCGGAGCTGTGAGAGCGCATGACAGAGGCTGGGCGACGCGGCGAGTGG	1080
Qy	1081	GCTGTTCATGTGCGCGCGCCGACGAGACGAGAGGTGTCAAGATCATGATGCTCTTC	1140
Db	1081	GCTGTTCATGTGCGCGCGCCGACGAGACGAGAGGTGTCAAGATCATGATGCTCTTC	1140
Qy	1141	AGAGCAGAGAACTGTGAGAGCAGATGCGCAGTGGCCGATTC	1200
Db	1141	AGAGCAGAGAACTGTGAGAGCAGATGCGCAGTGGCCGATTC	1200

1201 CGACGCTGACACAGCAGCGCATCAAGTTCAATGAAACGGGCTTATGGCCGACCCCAT 1260  
1201 CGACGCTGACACAGCAGCGCATCAAGTTCAATGAAACGGGCTTATGGCCGACCCCAT 1260  
1261 GAAAGGTGTACAAAGACCGCCAGGTCATGAAATGTGAGTGTAGCAGAGAGAGACCTT 1320  
1261 GAAAGGTGTACAAAGACCGCCAGGTCATGAAATGTGAGTGTAGCAGAGAGAGACCTT 1320  
1321 CCGGGGAAAGTTCATCAGATCCCAAGAACTTTGGCTGTGATCCATTCCTGAGAG 1380  
1321 CCGGGGAAAGTTCATCAGATCCCAAGAACTTTGGCTGTGATCCATTCCTGAGAG 1380  
1381 GAAAGACGTGCTGAGTGGCTCTCTATTACTACTGACTAAGAAATGAGAACTATTA 1440  
1381 GAAAGACGTGCTGAGTGGCTCTCTATTACTACTGACTAAGAAATGAGAACTATTA 1440  
1441 GAGCCTGTGTGAGACGGAGCTATCGGCGCCGCAAGAGCCAGCAGCAACACAGACGA 1500  
1441 GAGCCTGTGTGAGACGGAGCTATCGGCGCCGCAAGAGCCAGCAGCAACACAGACGA 1500  
1501 GCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
1501 GCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
1501 GCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
1561 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
1561 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
1621 CGACAAAGAGAGACTTCCTCAAGAGAGAGACAGACGACCTCAGGGAGAGACAAACGA 1680  
1621 CGACAAAGAGAGACTTCCTCAAGAGAGAGACAGACGACCTCAGGGAGAGACAAACGA 1680  
1681 GAAAGAGGCTGTGGCTCCAAAGGCGCAAAATGCCCAACAGCAGAGAGAGAGAGAG 1740  
1681 GAAAGAGGCTGTGGCTCCAAAGGCGCAAAATGCCCAACAGCAGAGAGAGAGAGAG 1740  
1741 CCGCATCACCCGCTCAATGAGCTAATGAGGCAACAGAGAGAGAGAGAGAGAGAGAG 1800  
1741 CCGCATCACCCGCTCAATGAGCTAATGAGGCAACAGAGAGAGAGAGAGAGAGAGAG 1800  
1801 GAGCGCGAGCTGGCTCCATGAGGCTGATGAGAGTTCCTGCGACAGAGAGAGAGAG 1860  
1801 GAGCGCGAGCTGGCTCCATGAGGCTGATGAGAGTTCCTGCGACAGAGAGAGAGAG 1860  
1861 GGAAGACGCGCAAGAGAGAGTCTCTGAGAACAGCGCCGCAACTGGTCGCGCATTC 1920  
1861 GGAAGACGCGCAAGAGAGAGTCTCTGAGAACAGCGCCGCAACTGGTCGCGCATTC 1920  
1921 GGTGGGCTCCAAAGACTGTGCGCAGTGTAAAGACTTCTAATTCAATCAAGAGAGAG 1980  
1921 GGTGGGCTCCAAAGACTGTGCGCAGTGTAAAGACTTCTAATTCAATCAAGAGAGAG 1980  
1981 GAACTCGATGATGATTTTCAAGCAGACAAAGCTGAAATGAGAGAGAGAGAGAGAG 2040  
1981 GAACTCGATGATGATTTTCAAGCAGACAAAGCTGAAATGAGAGAGAGAGAGAGAG 2040  
2041 GAGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
2041 GAGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
2101 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160  
2101 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160  
2161 GAGCTGAGCTTATCATGCTCTGTGGAGATGAGTGTCCAGAGAGAGAGAGAGAGAG 2220  
2161 GAGCTGAGCTTATCATGCTCTGTGGAGATGAGTGTCCAGAGAGAGAGAGAGAGAG 2220  
2221 GAGCTGAGCTTATCATGCTCTGTGGAGATGAGTGTCCAGAGAGAGAGAGAGAGAG 2280  
2221 GAGCTGAGCTTATCATGCTCTGTGGAGATGAGTGTCCAGAGAGAGAGAGAGAGAG 2280

2281 GGAACAAGGAGAGAGATGGGCTCCAGAGCCCGCAGCAGCTTGTGGCGCGAGCGGACAC 2340  
2281 GGAACAAGGAGAGAGATGGGCTCCAGAGCCCGCAGCAGCTTGTGGCGCGAGCGGACAC 2340  
2341 AGGCCCCACCCACCCACACAGAGAGACATCCCGGGGCCCCCATTTGAGACCCACCCCG 2400  
2341 AGGCCCCACCCACCCACACAGAGAGACATCCCGGGGCCCCCATTTGAGACCCACCCCG 2400  
2401 TGAAGGCAACCGAGAGCCCTACGCCCCCACAAGACCCCATCGCCTCTGACCTCTCC 2460  
2401 TGAAGGCAACCGAGAGCCCTACGCCCCCACAAGACCCCATCGCCTCTGACCTCTCC 2460  
2461 TGTGTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
2461 TGTGTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
2521 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
2521 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
2581 GCCCGTCAAGAGAGAGTGCACGGAGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
2581 GCCCGTCAAGAGAGAGTGCACGGAGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
2641 GAGAGCGCTGAGGCGCACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
2641 GAGAGCGCTGAGGCGCACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
2701 CCGCAGAGGCGCACCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760  
2701 CCGCAGAGGCGCACCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760  
2761 CTGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
2761 CTGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
2821 AAGGCGCAGAGCTCTTCAACCCCGACCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
2821 AAGGCGCAGAGCTCTTCAACCCCGACCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
2881 ACTGAGCTGAAAGCAGCTGAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
2881 ACTGAGCTGAAAGCAGCTGAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
2941 AGTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000  
2941 AGTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000  
3001 GCCACCGCAAAACCTGCAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
3001 GCCACCGCAAAACCTGCAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
3061 GGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120  
3061 GGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120  
3121 GAAAGCTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180  
3121 GAAAGCTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180  
3181 TGAAGTATCAAGAGCTTCCCGCAGATGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240  
3181 TGAAGTATCAAGAGCTTCCCGCAGATGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240  
3241 TGTGTACCCCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300  
3241 TGTGTACCCCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300  
3301 GAGCATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
3301 GAGCATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
3361 GCAAAATAGGTGCATCTCCCAAGAGAGATGTGGTCCAGCTTCCAGTCCGTAACAGAGA 3420



Db 3361 GCAAAATAGGATGCAATCTCCCAAGGAATGCGGTCCAGCTCCAGTCCCGTACTCCAGAGGA 3420  
OY 3421 TGCCAAAGGCCCCGGGTGGGCCCCCTGTCAACATGGAGGCTGCCCTGGCCATATGAAACCCAAAAA 3480  
Db 3421 TGCCAAAGGCCCCGGGTGGGCCCCCTGTCAACATGGAGGCTGCCCTGGCCATATGAAACCCAAAAA 3480  
OY 3481 GCTGGCACTCTTCAAGCGAGTAGAAGCAAGAGCAGCTGTCCCAAGGAGGAGGCTGGGCG 3540  
Db 3481 GCTGGCACTCTTCAAGCGAGTAGAAGCAAGAGCAGCTGTCCCAAGGAGGAGGCTGGGCG 3540  
OY 3541 ACCGGAAGACCTTGGGGGTGCCCCAGAGGCGTCCGTGCTGAAGAGGAGCAGCTT 3600  
Db 3541 ACCGGAAGACCTTGGGGGTGCCCCAGAGGCGTCCGTGCTGAAGAGGAGCAGCTT 3600  
OY 3601 GGGCTCAAGTTCGGGGCGGAAGATCAACAAAGGCAATTCACAGCAACAGGGTGCCTCGGA 3660  
Db 3601 GGGCTCAAGTTCGGGGCGGAAGATCAACAAAGGCAATTCACAGCAACAGGGTGCCTCGGA 3660  
OY 3661 CAGCGCCATCAATACCGGAGCTTCATCAACCAAGGCAAGGAGGAGGCTGATGATCTGTACAA 3720  
Db 3661 CAGCGCCATCAATACCGGAGCTTCATCAACCAAGGCAAGGAGGAGGCTGATGATCTGTACAA 3720  
OY 3721 GGGGCAACATCAACAGATATCATGGGCGAGAGACAGCCGAGTGGCTTGAACCGGGGCGGGA 3780  
Db 3721 GGGGCAACATCAACAGATATCATGGGCGAGAGACAGCCGAGTGGCTTGAACCGGGGCGGGA 3780  
OY 3781 GGAACAGCTTGCACAAAGGCGACAGTATCTACGAAGGCAAGAGGCGACAGTCTTGTCTTA 3840  
Db 3781 GGAACAGCTTGCACAAAGGCGACAGTATCTACGAAGGCAAGAGGCGACAGTCTTGTCTTA 3840  
OY 3841 TGAGGATGATCATGTCTGTGAACCAAGTGTCCAAAGAGGAGCGCAAGAGCAGCTCAGAGCC 3900  
Db 3841 TGAGGATGATCATGTCTGTGAACCAAGTGTCCAAAGAGGAGCGCAAGAGCAGCTCAGAGCC 3900  
OY 3901 CCCCCATGAGAGCGGCGCCCCCAAGCGCACTATGACATGATGAGAGGCGCGGTGGGCGAG 3960  
Db 3901 CCCCCATGAGAGCGGCGCCCCCAAGCGCACTATGACATGATGAGAGGCGCGGTGGGCGAG 3960  
OY 3961 AGCCATCTCTAGCAGAGTCAAGAGGCTCATGAGGCGTGCATCCGCGGAGCGACA 4020  
Db 3961 AGCCATCTCTAGCAGAGTCAAGAGGCTCATGAGGCGTGCATCCGCGGAGCGACA 4020  
OY 4021 CAGCCCCACCACTCAAAAGAGCAGACCAATCCGCGGCTCATCAACAAAGGATCCC 4080  
Db 4021 CAGCCCCACCACTCAAAAGAGCAGACCAATCCGCGGCTCATCAACAAAGGATCCC 4080  
OY 4081 TCGGTCCTACGTGAGAGCAAGAGAGGACTACCTGCGTGGAGGCGCAAGCTCTTAAAGCG 4140  
Db 4081 TCGGTCCTACGTGAGAGCAAGAGAGGACTACCTGCGTGGAGGCGCAAGCTCTTAAAGCG 4140  
OY 4141 GAGAGGCAAGCTCCGCCCCCAAGCGCTCAAGGAGCTGACCGAGGCTTACAAAGCGA 4200  
Db 4141 GAGAGGCAAGCTCCGCCCCCAAGCGCTCAAGGAGCTGACCGAGGCTTACAAAGCGA 4200  
OY 4201 GGGCCTTGGGCCCCCTGAAGCTGAAGCGGCGCATGAAGGCGCTGTGTGACCAAGGTAAGGA 4260  
Db 4201 GGGCCTTGGGCCCCCTGAAGCTGAAGCGGCGCATGAAGGCGCTGTGTGACCAAGGTAAGGA 4260  
OY 4261 GGGGAGGCGGCTTCATATGAGATCCCGGCGAGAGGCTGCGGCAACCGCGAAGCTGCG 4320  
Db 4261 GGGGAGGCGGCTTCATATGAGATCCCGGCGAGAGGCTGCGGCAACCGCGAAGCTGCG 4320  
OY 4321 CTTGGCCCCGCGGCTCAAGAGGAGCTCATCAAGAGGCGACCCGCTCAAGTACGA 4380  
Db 4321 CTTGGCCCCGCGGCTCAAGAGGAGCTCATCAAGAGGCGACCCGCTCAAGTACGA 4380  
OY 4381 CACCGGCGGCTCACCACTGGCTCCAAAAAGACAGACGTAAGCTCTCATGGGAGCCC 4440  
Db 4381 CACCGGCGGCTCACCACTGGCTCCAAAAAGACAGACGTAAGCTCTCATGGGAGCCC 4440  
OY 4441 CCGGCGGAGAGTTTCCACCCGTGCAACCGGTGATGTGATGAGCCAGCCCGGAGCACTGGA 4500  
Db 4441 CCGGCGGAGAGTTTCCACCCGTGCAACCGGTGATGTGATGAGCCAGCCCGGAGCACTGGA 4500

Db 4441 CCGGCGGAGAGTTTCCACCCGTGCAACCCGCTGATGTGATGAGCCAGCCCGGAGCACTGGA 4500  
OY 4501 AGTGCCTGTCTACAGAGAGAGCTTGAAGAGCGGCGCAAGAGACCGGCAAGCTCGGGGAG 4560  
Db 4501 AGTGCCTGTCTACAGAGAGAGCTTGAAGAGCGGCGCAAGAGACCGGCAAGCTCGGGGAG 4560  
OY 4561 CTCCATTGCGGCGGCGCGGCTCATTTGTGCTGAGCTGGGTAAAGCGGCGAGAGCCC 4620  
Db 4561 CTCCATTGCGGCGGCGCGGCTCATTTGTGCTGAGCTGGGTAAAGCGGCGAGAGCCC 4620  
OY 4621 CTTGACCTATAGAGACCAAGGGGACCCCTTTTGGGCGGACCCCTCCACAGAGGTTGGCCGT 4680  
Db 4621 CTTGACCTATAGAGACCAAGGGGACCCCTTTTGGGCGGACCCCTCCACAGAGGTTGGCCGT 4680  
OY 4681 GACCATGCGGAGAGCCCAAGCGGCGCTGACAGAGGAGGAGGAGGCTTGTGTCAGCAAGGATC 4740  
Db 4681 GACCATGCGGAGAGCCCAAGCGGCGCTGACAGAGGAGGAGGAGGCTTGTGTCAGCAAGGATC 4740  
OY 4741 CCAAGGACGAAAGCTGACGCTGAGAGCTGCTGTGAGATGCGCAAGTCCCGGACAGACCT 4800  
Db 4741 CCAAGGACGAAAGCTGAGAGCTGAGAGCTGCTGTGAGATGCGCAAGTCCCGGACAGACCT 4800  
OY 4801 GCGGAGGACCAACCCCACTGCGCCCTATGAGACCTGCTTGGGGGCTGAGTGG 4860  
Db 4801 GCGGAGGACCAACCCCACTGCGCCCTATGAGACCTGCTTGGGGGCTGAGTGG 4860  
OY 4861 GCGGAGGACCAACCCCACTGCGCCCTATGAGACCTGCTTGGGGGCTGAGTGG 4920  
Db 4861 GCGGAGGACCAACCCCACTGCGCCCTATGAGACCTGCTTGGGGGCTGAGTGG 4920  
OY 4921 CATCCCTTGGAGCGAGCGGCTGCTTACTTACCTGCGCGGACACCTTGGGCGGAGCCCAAC 4980  
Db 4921 CATCCCTTGGAGCGAGCGGCTGCTTACTTACCTGCGCGGACACCTTGGGCGGAGCCCAAC 4980  
OY 4981 CTACCCGACCTGTATGCCACCCCTACCTCATCGGCGGCTTACCCGAGCAAGGCGGCTGGA 5040  
Db 4981 CTACCCGACCTGTATGCCACCCCTACCTCATCGGCGGCTTACCCGAGCAAGGCGGCTGGA 5040  
OY 5041 GAAACCGGAGACCATCATCAATGAATGATCATGACCTTGGAGAGATGACCAACAGGAG 5100  
Db 5041 GAAACCGGAGACCATCATCAATGAATGATCATGACCTTGGAGAGATGACCAACAGGAG 5100  
OY 5101 CACCGCATGCGCCAGCGAGCTGATATGCTGAGGGGCTCTTGGCGCGGAGTCTTGGCT 5160  
Db 5101 CACCGCATGCGCCAGCGAGCTGATATGCTGAGGGGCTCTTGGCGCGGAGTCTTGGCT 5160  
OY 5161 GGCATCTCACTACGTTGAGGGGCCCCGAGGATCATGAGCCGTGCGCAAGTGCACACT 5220  
Db 5161 GGCATCTCACTACGTTGAGGGGCCCCGAGGATCATGAGCCGTGCGCAAGTGCACACT 5220  
OY 5221 GCTGTGTGCTGTGCCCCGAGACCAAGGACCCCAAGCGCACCGCATGAGACCGCTTGCCTA 5280  
Db 5221 GCTGTGTGCTGTGCCCCGAGACCAAGGACCCCAAGCGCACCGCATGAGACCGCTTGCCTA 5280  
OY 5281 CTTCCCAACCGGCGCCCAAGCCCTTACAGAGCGCGGCAAGAGCTTCCCATCTTCCCAAG 5340  
Db 5281 CTTCCCAACCGGCGCCCAAGCCCTTACAGAGCGCGGCAAGAGCTTCCCATCTTCCCAAG 5340  
OY 5341 AGGTCCAAACAATTGAACAAACCAACCAACGATCTGTGCGAGGCGGAGCGAGAGCG 5400  
Db 5341 AGGTCCAAACAATTGAACAAACCAACCAACGATCTGTGCGAGGCGGAGCGAGAGCG 5400  
OY 5401 GGAATCAGAGGCGGAGCCGGGATCGGAGACCGGAGAAAGTCCATCTTCAAGTCCACAGAC 5460  
Db 5401 GGAATCAGAGGCGGAGCCGGGATCGGAGACCGGAGAAAGTCCATCTTCAAGTCCACAGAC 5460  
OY 5461 GGTGAGACGACGACCTCATCTGAGAGCTGTGTACAGAGGAGGAGGAGGAGGAGGAG 5520  
Db 5461 GGTGAGACGACGACCTCATCTGAGAGCTGTGTACAGAGGAGGAGGAGGAGGAGGAG 5520  
OY 5521 CAGCGGCGGGGTGGGGGAGAGAGAGCGGCGGCTCCCATCTCCATGCGCAACAGCA 5580  
Db 5521 CAGCGGCGGGGTGGGGGAGAGAGAGCGGCGGCTCCCATCTCCATGCGCAACAGCA 5580

QY 5581 CTGGCCATCTCCCTCGAAGCCAGATGCGCTTCAGACAGAGACCCAGTGTGCTTCAAA 5640  
DB 5581 CTGGCCATCTCCCTCGAAGCCAGATGCGCTTCAGACAGAGACCCAGTGTGCTTCAAA 5640  
QY 5641 CACAGGCAATGAAGGATATCATCAACGCTGTGAGGCCCGAGAGCCCAACGCTCTGAAGTC 5700  
DB 5641 CACAGGCAATGAAGGATATCATCAACGCTGTGAGGCCCGAGAGCCCAACGCTCTGAAGTC 5700  
QY 5701 CACTTCACCTCTCTCAACCGCTTGGCCAGTGTGCAATTCCACCTGACCAACCACTGCCCC 5760  
DB 5701 CACTTCACCTCTCTCAACCGCTTGGCCAGTGTGCAATTCCACCTGACCAACCACTGCCCC 5760  
QY 5761 ACTGGGCGGCAACCTGATGGGGTCTACCTTACCTCATGGAAGCCGCTTGTGCGCCAA 5820  
DB 5761 ACTGGGCGGCAACCTGATGGGGTCTACCTTACCTCATGGAAGCCGCTTGTGCGCCAA 5820  
QY 5821 GGAAGCCCCCGGGGTGCCCCCGGCAAGGCGGCCCGAGAGACACCGGCGCATGCTTTCCT 5880  
DB 5821 GGAAGCCCCCGGGGTGCCCCCGGCAAGGCGGCCCGAGAGACACCGGCGCATGCTTTCCT 5880  
QY 5881 CGCCAAAGCCCCAGCCCGCTCCGGGCTGAGAGCCGCTCTCTCCCGAGCAAGGGCTCGGA 5940  
DB 5881 CGCCAAAGCCCCAGCCCGCTCCGGGCTGAGAGCCGCTCTCTCCCGAGCAAGGGCTCGGA 5940  
QY 5941 GCGCCGCGCCCTAGTGCCTCTGTCTGTGGCAAGCCACCATGCGCCGCAACCCCTGCGAA 6000  
DB 5941 GCGCCGCGCCCTAGTGCCTCTGTCTGTGGCAAGCCACCATGCGCCGCAACCCCTGCGAA 6000  
QY 6001 GAACCTCGACCTCAACCAAGCCAGCCCGGACCCGCGGCGCACCTGCTCGGCTCGGA 6060  
DB 6001 GAACCTCGACCTCAACCAAGCCAGCCCGGACCCGCGGCGCACCTGCTCGGCTCGGA 6060  
QY 6061 CCCGCAAGGGGAAAAGATCTCAAAAGTAAACCTTTTCAATCAAGAACTGGAACTCGTTTC 6120  
DB 6061 CCCGCAAGGGGAAAAGATCTCAAAAGTAAACCTTTTCAATCAAGAACTGGAACTCGTTTC 6120  
QY 6121 TCTGGGTTACACGAGAGAGCTACAGCCCGGAGGGGTGAGCCGCTAGGCCCTGTGAG 6180  
DB 6121 TCTGGGTTACACGAGAGAGCTACAGCCCGGAGGGGTGAGCCGCTAGGCCCTGTGAG 6180  
QY 6181 CTACCCCACTTGAACCCAGACAAAGGGGCTCCCAAGCACTTGAAGAGCTCGACAAGAG 6240  
DB 6181 CTACCCCACTTGAACCCAGACAAAGGGGCTCCCAAGCACTTGAAGAGCTCGACAAGAG 6240  
QY 6241 CCACTTGAAGGGGAGCTCGGCGCCAGCAAGCCAGCCCTGTGAGCTTGGCGGGAGGC 6300  
DB 6241 CCACTTGAAGGGGAGCTCGGCGCCAGCAAGCCAGCCCTGTGAGCTTGGCGGGAGGC 6300  
QY 6301 CGCCCACTCCCAACCTGCGGCGGCTGTGAGAGCCAGCCCTGTGAGAGCCCGGTGCT 6360  
DB 6301 CGCCCACTCCCAACCTGCGGCGGCTGTGAGAGCCAGCCCTGTGAGAGCCCGGTGCT 6360  
QY 6361 CCAAGACCGGCGGAGGGTCAAAAGTCAACAGCGGGGTGTCAACCTGTGAGCCAGCATAG 6420  
DB 6361 CCAAGACCGGCGGAGGGTCAAAAGTCAACAGCGGGGTGTCAACCTGTGAGCCAGCATAG 6420  
QY 6421 TGAAGTCAATCAACAGAGCTTACCCGAGCAACCAAGCAAGCTTCAAGGCAACCTGTGC 6480  
DB 6421 TGAAGTCAATCAACAGAGCTTACCCGAGCAACCAAGCAAGCTTCAAGGCAACCTGTGC 6480  
QY 6481 CGCCCCCTCTACTCTCTTCCCTGTGGGCGAGTGCCTCCGCTGTGAGACTTCCGCGCCAC 6540  
DB 6481 CGCCCCCTCTACTCTCTTCCCTGTGGGCGAGTGCCTCCGCTGTGAGACTTCCGCGCCAC 6540  
QY 6541 CAGTGAACCTTACTCTCCGCGCCCGGACCATGTGTGCGGCGCGCTCCCGCCACAG 6600  
DB 6541 CAGTGAACCTTACTCTCCGCGCCCGGACCATGTGTGCGGCGCGCTCCCGCCACAG 6600  
QY 6601 CGAAGGGGCGAAGAGTCTTCAGAGCCAAACAGAGCTGCGTCTTGGGTGTGTGAGGA 6660  
DB 6601 CGAAGGGGCGAAGAGTCTTCAGAGCCAAACAGAGCTGCGTCTTGGGTGTGTGAGGA 6660

QY 6661 CGGTATTGAACCTGTGTGCCCAACCGGAGGCAATGACGAGGCCAGGGCACTCCCGAGTGC 6720  
DB 6661 CGGTATTGAACCTGTGTGCCCAACCGGAGGCAATGACGAGGCCAGGGCACTCCCGAGTGC 6720  
QY 6721 TGTGTACCCGCTGTGTACCGGGATGGGAAACAGACGAGGCCAGAGATGGGCTTCA 6780  
DB 6721 TGTGTACCCGCTGTGTACCGGGATGGGAAACAGACGAGGCCAGAGATGGGCTTCA 6780  
QY 6781 GTCTCAGGCAACACAGCAGCGCGGCACTTCTTCAAGACCTACCGAAGCAATC 6840  
DB 6781 GTCTCAGGCAACACAGCAGCGCGGCACTTCTTCAAGACCTACCGAAGCAATC 6840  
QY 6841 CGCCATGTCTCAAGTCCAAAGAGCAAGATCAACAAGAACTGAACCAACCAACCGGAA 6900  
DB 6841 CGCCATGTCTCAAGTCCAAAGAGCAAGATCAACAAGAACTGAACCAACCAACCGGAA 6900  
QY 6901 TGAAGCTGAATACAAATATGAGCCAGCCGAGGAGCGAGATCTTCAATATGCCGATCAC 6960  
DB 6901 TGAAGCTGAATACAAATATGAGCCAGCCGAGGAGCGAGATCTTCAATATGCCGATCAC 6960  
QY 6961 CGGAAACAGGCTTTATGACCTTATGAAAGCCAGCGGTGTCAGAAACATGCCAGCAACAT 7020  
DB 6961 CGGAAACAGGCTTTATGACCTTATGAAAGCCAGCGGTGTCAGAAACATGCCAGCAACAT 7020  
QY 7021 GGGGCTGAGGCGCATTAATTAGAAAGGCACTAGGGTAAATATGACCAAGTGGGAAGTC 7080  
DB 7021 GGGGCTGAGGCGCATTAATTAGAAAGGCACTAGGGTAAATATGACCAAGTGGGAAGTC 7080  
QY 7081 CCGCGGCTCAAGCGCAATGCTTTTAAACCTCTGAATGCGAGTGCAGGCTGCGGTGC 7140  
DB 7081 CCGCGGCTCAAGCGCAATGCTTTTAAACCTCTGAATGCGAGTGCAGGCTGCGGTGC 7140  
QY 7141 TATGCCATTAACCGCTGTGAGAGGAGAGTACCAACATCTCACTCGCAGGTGGCGG 7200  
DB 7141 TATGCCATTAACCGCTGTGAGAGGAGAGTACCAACATCTCACTCGCAGGTGGCGG 7200  
QY 7201 CGGAAAGGCGCAAGGTCTGTGGAGAGCCAGACCCGAGCAAGCAAGTCCCGGCGCCG 7260  
DB 7201 CGGAAAGGCGCAAGGTCTGTGGAGAGCCAGACCCGAGCAAGCAAGTCCCGGCGCCG 7260  
QY 7261 CCTGGCATCTGGGAGCCGCGCAACCTCTGTCTCTCAAGTGCATCTCGAGGAGATGCA 7320  
DB 7261 CCTGGCATCTGGGAGCCGCGCAACCTCTGTCTCTCAAGTGCATCTCGAGGAGATGCA 7320  
QY 7321 CGCGGAGAGCGGCTTCAACCAACCGGTGTGGAGAGCAGGCCCTGCTCGCAGGTTCAC 7380  
DB 7321 CGCGGAGAGCGGCTTCAACCAACCGGTGTGGAGAGCAGGCCCTGCTCGCAGGTTCAC 7380  
QY 7381 GGCATTTCCCTTCAACCCCTGTATCATGCGGCTGACAGGCGGGGTGTGATGAGCTTCC 7440  
DB 7381 GGCATTTCCCTTCAACCCCTGTATCATGCGGCTGACAGGCGGGGTGTGATGAGCTTCC 7440  
QY 7441 CCCACCGGCGCTCCCGCGGAGGAGCGGCGCCCTGCTGCGCCCAACAGCCTGTGAGAC 7500  
DB 7441 CCCACCGGCGCTCCCGCGGAGGAGGCGGCGCCCTGCTGCGCCCAACAGCCTGTGAGAC 7500  
QY 7501 GGAAGCCAAAGCACTGTCTGTCTGCGAGTACAGAGCACTCTCGACAGAGAGTGTACG 7560  
DB 7501 GGAAGCCAAAGCACTGTCTGTCTGCGAGTACAGAGCACTCTCGACAGAGAGTGTACG 7560  
QY 7561 AACAAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 7620  
DB 7561 AACAAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 7620  
QY 7621 CGGCGCTGAGAGAGCGGAGCGGCTGCGGCACTCCCAACCAAGAGAGAGCCCTGAGTC 7680  
DB 7621 CGGCGCTGAGAGAGCGGAGCGGCTGCGGCACTCCCAACCAAGAGAGAGCCCTGAGTC 7680  
QY 7681 CGGCGCTGAGAGAGCGGAGCGGCTGCGGCACTCCCAACCAAGAGAGAGCCCTGAGTC 7740  
DB 7681 CGGCGCTGAGAGAGCGGAGCGGCTGCGGCACTCCCAACCAAGAGAGAGCCCTGAGTC 7740  
QY 7741 AACTAAGACTCCGCGCGGAGGCTGCGCTGTGACAGACTTACTAGGGGATGTTTACTG 7800

Db	7731	AACCTBAAGACTCCCGCCCGGGGCTGGCCCTGTGTGCAGACTTACTCAGGGGAGATGTTTACTCTG	7790
QY	7801	GTEGCTCGGGAAGGAGGAGGGAAGGGGCGCGGGAGGGGGGACGCGACGGCGTGTGGCAGCCAC	7860
Db	7791	GTEGCTCGGGAAGGAGGAGGGAAGGGGCGCGGGAGGGGGGACGCGACGGCGTGTGGCAGCCAC	7850
QY	7861	ACACAGGCGGCGCCAGGGCGGGCCAGGGAGCCCAAAAGCAGAGTGAACACACGACTCTTCACGCCAC	7920
Db	7851	ACACAGGCGGCGCCAGGGCGGGCCAGGGAGCCCAAAAGCAGAGTGAACACACGACTCTTCACGCCAC	7910
QY	7921	TGCGCTCCCCGAGATGCAATTTGGAACCAAAAGCTTAACTGAGCTGCGACGCCCGCGGCCCT	7980
Db	7911	TGCGCTCCCCGAGATGCAATTTGGAACCAAAAGCTTAACTGAGCTGCGACGCCCGCGGCCCT	7970
QY	7981	CCCTCCGCGCTTCCCAATCCGCTTTAGCGCTCTTGAGACAGATGACGACGAGGCCCTGTGCCAGCC	8040
Db	7971	CCCTCCGCGCTTCCCAATCCGCTTTAGCGCTCTTGAGACAGATGACGACGAGGCCCTGTGCCAGCC	8030
QY	8041	CCAGTGGCGCTCGTTCCGGTCCCGCACAGACTGCCCGCACGACGAGATTTGCTGGAACCA	8100
Db	8031	CCAGTGGCGCTCGTTCCGGTCCCGCACAGACTGCCCGCACGAGATTTGCTGGAACCA	8090
QY	8101	GTCAGGCGCAGGTGGGGCGGACAAAAGGCGCAGGTGCGGCTGGGGGGGAACGGAATGCTCCGA	8160
Db	8091	GTCAGGCGCAGGTGGGGCGGACAAAAGGCGCAGGTGCGGCTGGGGGGGAACGGAATGCTCCGA	8150
QY	8161	GGACTGGACTGTTTTTTTTTCACACATCGTTTGGCGACGCGGTGGGAGAAAGGACAGATGA	8220
Db	8151	GGACTGGACTGTTTTTTTTTCACACATCGTTTGGCGACGCGGTGGGAGAAAGGACAGATGA	8210
QY	8221	AATGATGTGTGTGTTTACAGGGGTATATTTTGTATCCTTCATGATATTAATTCAGATGTT	8280
Db	8211	AATGATGTGTGTGTTTACAGGGGTATATTTTGTATCCTTCATGATATTAATTCAGATGTT	8270
QY	8281	TTACGCAAGGAAGGACTTACCAGATTTACTGTGCTGTGTGCTTTTGATCTGTGCTTACCG	8340
Db	8271	TTACGCAAGGAAGGACTTACCAGATTTACTGTGCTGTGTGCTTTTGATCTGTGCTTACCG	8330
QY	8341	TTTCAAGAGGCGGTGTGCAGGCGCGGACAGTGCCTGACCCCATCACTGCGAGAGCAAAAGGGGCG	8400
Db	8331	TTTCAAGAGGCGGTGTGTGCAGGCGCGGACAGTGCCTGACCCCATCACTGCGAGAGCAAAAGGGGCG	8390
QY	8401	GGGAGACTGCTGTCACGCGCCGCGTGTGCTCCCTCCCTCCCTTCCTTTGGGACAGATGA	8460
Db	8391	GGGAGACTGCTGTCACGCGCCGCGTGTGCTCCCTCCCTCCCTTCCTTTGGGACAGATGA	8450
QY	8461	TTTGCATGCGTATTCGTGTGGCGCGCAATTTGCGCAGAGGTGTGTATTCCTGATTTACCA	8520
Db	8451	TTTGCATGCGTATTCGTGTGGCGCGCAATTTGCGCAGAGGTGTGTATTCCTGATTTACCA	8510
QY	8521	CGTGGCTTATATTTAAAGGCAATTTATCTCCAAAAAATTTTTTTTTTTTTTTTTTTTTT	8571
Db	8511	CGTGGCTTATATTTAAAGGCAATTTATCTCCAAAAAATTTTTTTTTTTTTTTTTTTTTT	8561

RESULT 3	
AAC74783	
ID	AAC74783 standard; cDNA; 8564 BP.
XX	
AC	AAC74783;
XX	
DT	08-FEB-2001 (first entry)
XX	
DE	Human ORFX ORF38 polynucleotide sequence SEQ ID NO:675.
XX	
KW	Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW	vulnery; antipsoatic; antiparkinsonian; nootropic; neuroprotective;
KW	anticoagulant; osteopathic; antiarthritic; immunosuppressant; cardiact;
KW	immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW	hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW	antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW	antianaemic; gene therapy; cancer; proliferative disorder; hypertension;

KW	neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW	cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW	cholesterol ester storage; systemic lupus erythematosus; infection;
KW	severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW	allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW	bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW	thrombosis; contraceptive; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200058473-A2.
XX	
PD	05-OCT-2000.
XX	
PF	31-MAR-2000; 2000WO-US008621.
XX	
PR	31-MAR-1999; 99US-0127607P.
PR	02-APR-1999; 99US-0127636P.
PR	05-APR-1999; 99US-0127728P.
PR	30-MAR-2000; 2000US-00540763.
XX	
PA	(CUTRA-) CUTRAGEN CORP.
XX	
PI	Shinkets RA, Leach M;
XX	
DR	WPI: 2000-602362/57.
XX	
DR	P-PSDB; AAB40574.
XX	
PT	Novel nucleic acids and peptides derived from open reading frame X,
PT	useful for treating e.g. cancers, proliferative disorders,
PT	neurodegenerative disorders and cardiovascular disease.
XX	
SS	Claim 5; Page 779-784; 5507bp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cytostatic; hepatotropic; vlnary;  
CC antiporatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
CC anticonvulsant; antirheumatic; immunosuppressant; immunostimulant;  
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
CC antiviral; antifungal; antirheumatic; antihyroid; and antinaemic. The  
CC sequences can be used for determining the presence of or predisposition  
CC to, or preventing or treating pathological conditions associated with an  
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
CC used to treat cancers, proliferative disorders, neurodegenerative  
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease  
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease, to  
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
XQ  
XQ Sequence 8564 BP; 1859 A; 3034 C; 2532 G; 1139 T; 0 U; 0 Other;

Sequence 8564 BP; 1859 A; 3034 C; 2532 G; 1139 T; 0 U; 0 Other;

Query Match	99.2%	Score 8495.2;	DB 3;	Length 8564;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 8546; Conservative	1;	Mismatches 13;	Indels 14;	Gaps 3

Qy	2	ATGTGGGGCTCCACACAGTTGTGGCACAGAGTGGAAAGGCACTAGACCCCGTTACCG	61
Db	1	ATGTGGGGCTCCACACAGTTGTGGCACAGAGTGGAAAGGCACTAGACCCCGTTACCG	60
Qy	62	CCCCAAGGCTTTCTACCCAGTAGATCGCCGGAGCGCACAGGAGTTCGGGCTCTG	122
Db	61	CCCCAAGGCTTTCTACCCAGTAGAGATCGCCGGAGCGCACAGGAGTTCGGGCTCTG	120
Qy	122	GAGTACACACACACATCCCGGAGCTATGCTCCACCTGTGCCGGGCTCATCATCCAG	181
Db	121	GAGTACACACACACATCCCGGAGCTATGCTCCACCTGTGCCGGGCTCATCATCCAG	180

OY	182	CCCCCAGCGCGAGAGGCCCTCCCTGCTGTCTAGATTCCAGCCCGGAAATGAACGGTCCAG	241
Db	181	CCCCAGCGCGAGAGGCCCTCCCTGCTGTCTAGATTCCAGCCCGGAAATGAACGGTCCAG	240
OY	242	GAGCTTCCACTGCGCGCCAGAGATCCCACTCATCTGCGCCGAGCTGCGGAAATGTCAGAGATG	301
Db	241	GAGCTTCCACTGCGCGCCAGAGATCCCACTCATCTGCGCCGAGCTGCGGAAATGTCAGAGATG	300
OY	302	GAGTTCAATTGAAAGCAAGCGCCCTCGGCTTAGAGCTGCGCTGACCCCTGCTGACCG	361
Db	301	GAGTTCAATTGAAAGCAAGCGCCCTCGGCTTAGAGCTGCGCTGACCCCTGCTGACCG	360
OY	362	TCACCCTCTGCTGCGCCAGCGCCAGCTTGCGGATTTGAAAGCTTACCAAGACCTGTAGC	421
Db	361	TCACCCTCTGCTGCGCCAGCGCCAGCTTGCGGATTTGAAAGCTTACCAAGACCTGTAGC	420
OY	422	CTGACGGGCAAGCTGAGAAACCGGTGTCTCCCCCAGCCCCCGGACATGACCTTGAGCTG	481
Db	421	CTGACGGGCAAGCTGAGAAACCGGTGTCTCCCCCAGCCCCCGGACATGACCTTGAGCTG	480
OY	482	GAGCTGTGTCCGCCACCGCTGTCCAAAGAGAGCTGATCCAGAACATGSAACCGCGTGCAC	541
Db	481	GAGCTGTGTCCGCCACCGCTGTCCAAAGAGAGCTGATCCAGAACATGSAACCGCGTGCAC	540
OY	542	CGAGAGATCAACATGTAGAGCAAGATCTTTAACTGTAAGAGAGAACACAAACAGCTG	601
Db	541	CGAGAGATCAACATGTAGAGCAAGATCTTTAACTGTAAGAGAGAACACAAACAGCTG	600
OY	602	GAGAGGAGGCGTGGCCAAACCCCGCGAGCTGAGAAACCCCGTGTACCGCGCCCATGCGAC	661
Db	601	GAGAGGAGGCGTGGCCAAACCCCGCGAGCTGAGAAACCCCGTGTACCGCGCCCATGCGAC	660
OY	662	TGGAAGCACCGGAGCCTGCTGTGAGATCATCTTACGACGAGAAACCGGAAAGAGCTGAAGCT	721
Db	661	TGGAAGCACCGGAGCCTGCTGTGAGATCATCTTACGACGAGAAACCGGAAAGAGCTGAAGCT	720
OY	722	GCACATCGGATTTCTGAAAGGCTTGGGGGCCCAAGTGAGCTGCGCTGTACCAACCAAGCC	781
Db	721	GCACATCGGATTTCTGAAAGGCTTGGGGGCCCAAGTGAGCTGCGCTGTACCAACCAAGCC	780
OY	782	TCCGACACCCCGCAGTATCATGAGAAACATCAAAATTAACCAAGCGGATGCGGAAAGAGCTA	841
Db	781	TCCGACACCCCGCAGTATCATGAGAAACATCAAAATTAACCAAGCGGATGCGGAAAGAGCTA	840
OY	842	ATCTTGTACTTCAAGAGAGGAAATCAACCTCGGAAACATGGAAGCAAAATTCTGCGAG	901
Db	841	ATCTTGTACTTCAAGAGAGGAAATCAACCTCGGAAACATGGAAGCAAAATTCTGCGAG	900
OY	902	CGCTATGACCAAGCTCATGSAAGGCTTTGGAAAAAAAGTGAAGCGCATGAAAACAACCGC	961
Db	901	CGCTATGACCAAGCTCATGSAAGGCTTTGGAAAAAAAGTGAAGCGCATGAAAACAACCGC	960
OY	962	CGCGCGCGCGGCCAAGAGAGCAAGGTGGCCGAGTCTACGAAAAAGCAATTCCCTGAGATC	1021
Db	961	CGCGCGCGCGGCCAAGAGAGCAAGGTGGCCGAGTCTACGAAAAAGCAATTCCCTGAGATC	1020
OY	1022	CGCAAGCAGCGGAGCTGAGGAGGCGATGCAAGACAGAGGTTGGGCGACGCGGGGACATGCGG	1081
Db	1021	CGCAAGCAGCGGAGCTGAGGAGGCGATGCAAGACAGAGGTTGGGCGACGCGGGGACATGCGG	1080
OY	1082	CTGTTCATGTGCGCGCGCCGACGAGCAAGAGGTGTCAAGATCATGATGCGCTCTCA	1141
Db	1081	CTGTTCATGTGCGCGCGCCGACGAGCAAGAGGTGTCAAGATCATGATGCGCTCTCA	1140
OY	1142	GAGCAGAGAAACTTGGAGAGCAGATGCGCCAGCTGCGCTGTGATCCGCGCATGTGTAGC	1201
Db	1141	GAGCAGAGAAACTTGGAGAGCAGATGCGCCAGCTGCGCTGTGATCCGCGCATGTGTAGC	1200
OY	1202	GACGCTGACCAAGCAGCGGATCAAGTTCAATCAACATGAAACGGGCTTATGGCGGACCCCATG	1261
Db	1201	GACGCTGACCAAGCAGCGGATCAAGTTCAATCAACATGAAACGGGCTTATGGCGGACCCCATG	1260
OY	1262	AAGGTGTACAAAGACCGCAGGCTCATGACAATGTGAAGTGAAGAGAAAGAGACCTTC	1321

[illegible]

Db 2341 CAGAGCCACCCACCGAGACATCCCGGCCCCCATTTAGCCACCCGGCC 2400  
Qy TCTGAAGCCCGAGGCCCTTACGCCCCCAGACACCCCAATGCGCTTGTCACTCT 2458  
Db TCTGAAGCCCGAGGCCCTTACGCCCCCAGACACCCCAATGCGCTTGTCACTCT 2460  
Qy CCTGTGTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2518  
Db CCTGTGTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
Qy GGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2578  
Db GGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
Qy GAGCCCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2638  
Db GAGCCCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
Qy GCGAGAGCGCTGAGAGCCAGGCGGAGGCGCTTCAAGGCAAGAGAGAGAGAGAG 2698  
Db GCGAGAGCGCTGAGAGCCAGGCGGAGGCGCTTCAAGGCAAGAGAGAGAGAGAG 2700  
Qy AGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2758  
Db AGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760  
Qy ACCTGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2818  
Db ACCTGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
Qy CCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2878  
Db CCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
Qy CCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2938  
Db CCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
Qy AAAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2998  
Db AAAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000  
Qy CCGGCAACCGCAAAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3058  
Db CCGGCAACCGCAAAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
Qy CCGGCAACCGCAAAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3118  
Db CCGGCAACCGCAAAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120  
Qy CAGAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3178  
Db CAGAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180  
Qy CGTGAAGTGAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3238  
Db CGTGAAGTGAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240  
Qy CCGTGAACCGCAAAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3298  
Db CCGTGAACCGCAAAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300  
Qy CCGCAACATCTCAAAACCGGCTTCCCTCAATCTCTCTGCAAGAGAGAGAGAGAG 3358  
Db CCGCAACATCTCAAAACCGGCTTCCCTCAATCTCTCTGCAAGAGAGAGAGAGAG 3360  
Qy AGGCAAAATAGTGCATCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3418  
Db AGGCAAAATAGTGCATCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420  
Qy CATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3478  
Db CATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3480

Qy AAGCTGAGACCTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3479  
Db AAGCTGAGACCTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3481  
Qy CCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3538  
Db CCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540  
Qy CCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3598  
Db CCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600  
Qy CCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3658  
Db CCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3660  
Qy GACAGCGCATCATCAATACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3718  
Db GACAGCGCATCATCAATACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720  
Qy AAGGCAACATCAACAGAGATCATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3778  
Db AAGGCAACATCAACAGAGATCATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780  
Qy GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3838  
Db GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3840  
Qy TATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3898  
Db TATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3900  
Qy CCCCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3958  
Db CCCCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3960  
Qy AAGGCAATCTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4018  
Db AAGGCAATCTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4020  
Qy CACAGCCCCCAACCACTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4078  
Db CACAGCCCCCAACCACTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4080  
Qy CCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4138  
Db CCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4140  
Qy CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4198  
Db CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4200  
Qy CAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4258  
Db CAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4260  
Qy GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4318  
Db GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4320  
Qy CCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4378  
Db CCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4380  
Qy GACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4438  
Db GACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4440  
Qy CCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4498  
Db CCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4500  
Qy GAAAGTGCCTGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4558  
Db GAAAGTGCCTGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560



4559 GGGTCATTGCGCGCGCGCCCGCTCATTTGCTGAGGTGGGTAAAGCCGCGCAGAGC 4618  
4561 GGGTCATTGCGCGCGCGCCCGCTCATTTGCTGAGGTGGGTAAAGCCGCGCAGAGC 4620  
4619 CCCCTGACCTTAGAGACCAACGCGGGCAACCTTTGGCGGCACTCCCAAGAGTTGGCC 4678  
4621 CCCCTAACCTATAGAGACCAACGCGGGCAACCTTTGGCGGCACTCCCAAGAGTTGGCC 4680  
4679 GTGACCATGGGGAGGCCACCGCGCGCTGAGAGAGGGGCAAGCTTTGCTCAAGAGGCA 4738  
4681 GTGACCAACGGGGAGGCCACCGCGCGCTGAGAGAGGGGCAAGCTTTGCTCAAGAGGCA 4740  
4739 TCCCAAGACCGAAAGGTGAGCTGAGAGCTGTGAGATGCGCAAGTCCCGCAACAGACC 4798  
4741 TCCCAAGACCGAAAGGTGAGCTGAGAGCTGTGAGATGCGCAAGTCCCGCAACAGACC 4800  
4799 GTGCGCGAGACCAACCAACCCCATCTGCGCTTAGAGCACTTGTTGGGGCTGAGT 4858  
4801 GTGCGCGAGACCAACCAACCCCATCTGCGCTTAGAGCACTTGTTGGGGCTGAGT 4860  
4859 GGGGTGACCTGTATGCGAGGCAATCCCGCTGGGCTTGACCCCACTTCATACCCCGC 4918  
4861 GGGGTGACCTGTATGCGAGGCAATCCCGCTGGGCTTGACCCCACTTCATACCCCGC 4920  
4919 GGCATCCTCTGAGACGACGCGCTGCTACTTACCTGCGCCGACACCTGCGCCCAACCC 4978  
4921 GGCATCCTCTGAGACGACGCGCTGCTACTTACCTGCGCCGACACCTGCGCCCAACCC 4980  
4979 ACCTACCCGCACTGTATCCCAACCTTACCTTATCAGCGGCTTACCCGACACGAGCGGCTG 5038  
4981 ACCTACCCGCACTGTATCCCAACCTTACCTTATCAGCGGCTTACCCGACACGAGCGGCTG 5040  
5039 GAGAACCGGACGACATCATCATATGACTATACCTTCGAGAGATGACCAACAACG 5098  
5041 GAGAACCGGACGACATCATCATATGACTATACCTTCGAGAGATGACCAACAACG 5100  
5099 GCCACCGGCACTGACGAGGCTGATATGCTGAGGGGCTCTGCGCCCGGAGTCCG 5158  
5101 GCCACCGGCACTGACGAGGCTGATATGCTGAGGGGCTCTGCGCCCGGAGTCCG 5160  
5159 CTGGCACTCAACTAGCTGCGGGTCCCGAGGACATGACCTGTTCGAAGTCCAC 5218  
5161 CTGGCACTCAACTAGCTGCGGGTCCCGAGGACATGACCTGTTCGAAGTCCAC 5220  
5219 CTGCTGTGTCTGTGCTGCGCGGACCAAGGACCCCGCACTGATGACCCGCTTGGC 5278  
5221 CTGCTGTGTCTGTGCTGCGCGGACCAAGGACCCCGCACTGATGACCCGCTTGGC 5280  
5279 TACCTCCCAACCGCGCCGCTTCAAGAGCGGCAAGAGCTCCCACTTCCCA 5338  
5281 TACCTCCCAACCGCGCCGCTTCAAGAGCGGCAAGAGCTCCCACTTCCCA 5340  
5339 GAGAGTCAACACACTTGAACAAACCAACCAACGCTCTCTGTCGAGCGGAGCGAGAC 5398  
5341 GAGAGTCAACACACTTGAACAAACCAACCAACGCTCTCTGTCGAGCGGAGCGAGAC 5400  
5399 CGGGATTCGAGAGCGGACCGGGATTCGGAGCGGGAAAAATTCATCTCTCACTCAACG 5458  
5401 CGGGATTCGAGAGCGGACCGGGATTCGGAGCGGGAAAAATTCATCTCTCACTCAACG 5460  
5459 ACAGTGAAGCAGCAACCATCTGAGACCTGTATCAAGAGAGAGAGCGGCAAGCGC 5518  
5461 ACAGTGAAGCAGCAACCATCTGTGAGCTGTATCAAGAGAGAGAGCGGCAAGCGC 5520  
5519 AGCAGCGCGGGGTGGGGCAGACAGAGCGCGCGCTTCCCACTTCCATGCGCACG 5578  
5521 AGCAGCGCGGGGTGGGGCAGACAGAGCGCGCGCTTCCCACTTCCATGCGCACG 5580  
5579 CATCTGCGCATCTTCTCTGAGACCCAGAGATGCTTCAAGAGAGACCAATGTGCTTAC 5638  
5581 CATCTGCGCATCTTCTCTGAGACCCAGAGATGCTTCAAGAGAGACCAATGTGCTTAC 5640  
5639 AACACAGCATGAAGGTATCATCACCGTGTGAGGCCGACGAAGCCACGCTCTGAGG 5698

5641 AACACAGCATGAAGGTATCATCACCGTGTGAGGCCGACGCGCATCGTCTGAGG 5700  
5699 TCACCTCAACCTTCAACCCGCTGCGCCAGCGTGCACANTTCCACCTGCAACCACTG 5758  
5701 TCACCTCAACCTTCAACCCGCTGCGCCAGCGTGCACANTTCCACCTGCAACCACTG 5760  
5759 CCACTGGGCGCAACCTTCATAGGGGTCTACCTTACCTTCAATGAGGCCGTCTTGTGCCC 5818  
5761 CCACTGGGCGCAACCTTCATAGGGGTCTACCTTACCTTCAATGAGGCCGTCTTGTGCCC 5820  
5819 AAGAGAGCCCCCGGGTCCCGCGCCAGAGCGGCCCGGACAGACACCGGCAATGCTTTC 5878  
5821 AAGAGAGCCCCCGGGTCCCGCGCCAGAGCGGCCCGGACAGACACCGGCAATGCTTTC 5880  
5879 CTGCGCAACCCCGACCCGCTCCGGGCTGAGAGCCGCTTCTCCCGCAAGAGGCTCG 5938  
5881 CTGCGCAACCCCGACCCGCTCCGGGCTGAGAGCCGCTTCTCCCGCAAGAGGCTCG 5940  
5939 GAGCCCGGCGCTTATGCTCTCTGCTCTGAGCAGCAACATCGACCGCAACCCCTGCG 5998  
5941 GAGCCCGGCGCTTATGCTCTCTGCTCTGAGCAGCAACATCGACCGCAACCCCTGCG 6000  
5999 AAGAACCTGCACTTCAACGCGCAGCGCCGAGCCGCGCGCACCTGCTGCGCTCG 6058  
6001 AAGAACCTGCACTTCAACGCGCAGCGCCGAGCCGCGCGCACCTGCTGCGCTCG 6060  
6059 GACCGGCAACGGGAAAAAGCTCAAAATTAACCTTTTCCATCCAGAACTGGAATCCGT 6118  
6061 GACCGGCAACGGGAAAAAGCTCAAAATTAACCTTTTCCATCCAGAACTGGAATCCGT 6120  
6119 TCTCTGGATTACACGCGCAGCAGCTTACAGCCCGGAGAGGGTGAAGCCGTCAAGCTGTG 6178  
6121 TCTCTGGATTACACGCGCAGCAGCTTACAGCCCGGAGAGGGTGAAGCCGTCAAGCTGTG 6180  
6179 AGCTCAACCAAGTGTGACCAACGACAAAGGGCTTCCCAAGCACTTGAAGTCTGACAG 6238  
6181 AGCTCAACCAAGTGTGACCAACGACAAAGGGCTTCCCAAGCACTTGAAGTCTGACAG 6240  
6239 AGCCACTTGAAGGGGAGAGCTGCGGCCCAAGACCAAGGCCCGCTTGAAGCTTGGCGGGAG 6298  
6241 AGCCACTTGAAGGGGAGAGCTGCGGCCCAAGACCAAGGCCCGCTTGAAGCTTGGCGGGAG 6300  
6299 GCGGCCAACCTCCCAACCTGCGCGCGCTGCTGAGAGCAGCGCTCTGTCAGCCCGCTG 6358  
6301 GCGGCCAACCTCCCAACCTGCGCGCGCTGCTGAGAGCAGCGCTCTGTCAGCCCGCTG 6360  
6359 CTCAGACCGCCCAAGGGTCAAAAGGTCAACAGCGGGTGTCACTTGGCTCAGCAATC 6418  
6361 CTCAGACCGCCCAAGGGTCAAAAGGTCAACAGCGGGTGTCACTTGGCTCAGCAATC 6420  
6419 AGTGAAGTATCAACAGAGCTTACACCCCGGACCAACCAAGAGAGCTCAAGCGACCTTG 6478  
6421 AGTGAAGTATCAACAGAGCTTACACCCCGGACCAACCAAGAGAGCTCAAGCGACCTTG 6480  
6479 CCGGCCCGCTTCACTCTTCTCTGAGGCAAGTGTGCTTCTGAGCTTCCGCGGCCCA 6538  
6481 CCGGCCCGCTTCACTCTTCTCTGAGGCAAGTGTGCTTCTGAGCTTCCGCGGCCCA 6540  
6539 CCCAGTGAACCTTCACTTCCCGCCCGGACCAATGATGCGCGCGCTTCCCGCAC 6598  
6541 CCCAGTGAACCTTCACTTCCCGCCCGGACCAATGATGCGCGCGCTTCCCGCAC 6600  
6599 AGCGAAGGGGCAAGAGTCTTCAAGAGCCAAACAAAGAGCTGCTTGGGTGGTGGAG 6658  
6601 AGCGAAGGGGCAAGAGTCTTCAAGAGCCAAACAAAGAGCTGCTTGGGTGGTGGAG 6660  
6659 GACGATTTGAACCTGATGCCCAACCGAGGGGATGACGAGCCAGGCACTCCGGAAT 6718  
6661 GACGATTTGAACCTGATGCCCAACCGAGGGGATGACGAGCCAGGCACTCCGGAAT 6720  
6719 GCTGTATCCCGTGTGTATCCGGATGAGGAAACAGCGAGCCACGAGATGGGCTTC 6778

Db 6721 GCTGTGTACCCGCTGCTGTACCGGGATGGGGAAACAGACGAGCCACGACGAGATGGGCTCC 6780  
OY AAGTCTCCAGGCAACACGAGCCAGCGCCAGGCTTCTTTCAGCAAGGTGACCGAGAGCAAC 6838  
Db 6781 AAGTCTCCAGGCAACACGAGCCAGCGCCAGGCTTCTTTCAGCAAGGTGACCGAGAGCAAC 6840  
OY TCCGCGCATGTGTAAGTCCAGAGAGAGAGATCAACAAAGAGCTGAAACCCCAACCGG 6898  
Db 6841 TCCGCGCATGTGTAAGTCCAGAGAGAGAGATCAACAAAGAGCTGAAACCCCAACCGG 6900  
OY AATGAGCCCTGAATATCAATATCAAGCCAGCTGGAGACGGAATCTTCAATATGCCCCGATC 6958  
Db 6901 AATGAGCCCTGAATATCAATATCAAGCCAGCTGGAGACGGAATCTTCAATATGCCCCGATC 6960  
OY ACCGGAACAGGCGCTTATGACTATAGAAAGCAGGCGGTGACAGAAACATSCAGACCAAC 7018  
Db 6961 ACCGGAACAGGCGCTTATGACTATAGAAAGCAGGCGGTGACAGAAACATSCAGACCAAC 7020  
OY ATGGGGCTGGAGGCAATTAATGAAAGGCACTCATGGGTAAATATATGACCAATGGGAAAG 7078  
Db 7021 ATGGGGCTGGAGGCAATTAATGAAAGGCACTCATGGGTAAATATATGACCAATGGGAAAG 7080  
OY TCCCGCGCGCTCAGGCGCAATGCTTTTAAACCTCTGAAATGCCAGTGCACGCTCGCCGCT 7138  
Db 7081 TCCCGCGCGCTCAGGCGCAATGCTTTTAAACCTCTGAAATGCCAGTGCACGCTCGCCGCT 7140  
OY GCTATGACCATAAACCCTGCTGACGAGACGAGATGACCAACACTCACTCGCCAGGTGC 7198  
Db 7141 GCTATGACCATAAACCCTGCTGACGAGACGAGATGACCAACACTCACTCGCCAGGTGC 7200  
OY GCGGGGAAAGGCGCAAGGTCTCTGGCAGACCCAGACGCGGAAAGGCAAGTCCCGGCGCG 7258  
Db 7201 GCGGGGAAAGGCGCAAGGTCTCTGGCAGACCCAGACGCGGAAAGGCAAGTCCCGGCGCG 7260  
OY GGCCTGCGATCTGGGGGACCGGCGCAACCTTGTCTCTCACTGACGACCTCGGAGGGAATGCG 7318  
Db 7261 GGCCTGCGATCTGGGGGACCGGCGCAACCTTGTCTCTCACTGACGACCTCGGAGGGAATGCG 7320  
OY AACCCGCGGACGCGCTCAACACCGCGGTGTGGAGAGACAGGCCCTCGTCCGACGATTC 7378  
Db 7321 AACCCGCGGACGCGCTCAACACCGCGGTGTGGAGAGACAGGCCCTCGTCCGACGATTC 7380  
OY ACGCATTCCTCCCTACCAACCCCTGATCATGCGGCTGACAGCGGGGTGATGCTTCCCA 7438  
Db 7381 ACGCATTCCTCCCTACCAACCCCTGATCATGCGGCTGACAGCGGGGTGATGCTTCCCA 7440  
OY CCCCCACCGGGCTCTCCCGCGGCGAGCGGCGCCCTTGTGCGCCGCCCAACGCGCTGGGAC 7498  
Db 7441 CCCCCACCGGGCTCTCCCGCGGCGAGCGGCGCCCTTGTGCGCCGCCCAACGCGCTGGGAC 7500  
OY GAGAGACCCCAACCTGCTGTGCTGACGATACGAGACACTCTCCGACAGCGAGTGAATC 7558  
Db 7501 GAGAGACCCCAACCTGCTGTGCTGACGATACGAGACACTCTCCGACAGCGAGTGAATC 7560  
OY AGAACAGGCGCGGGCGGGCGGGCGGTGEBERS PERLUGTCAAGTCCCAAGCAAGCAAG 7618  
Db 7561 AGAACAGGCGCGGGCGGGCGGGCGGGCGG-----TGTCAAGTCCCAAGCAAGCAAG 7610  
OY AACGCGCTCTGACGAGCGGGCGGGCTGCGGACTCCCCCAACCAAGAAAGAACCCCTGAG 7678  
Db 7611 AACGCGCTCTGACGAGCGGGCGGGCTGCGGACTCCCCCAACCAAGAAAGAACCCCTGAG 7670  
OY TCCGCGCTGCGCTCATCATCTGTCCGTCAGAGCGCGGACATCTTGCCTGTAAAGCC 7738  
Db 7671 TCCGCGCTGCGCTCATCATCTGTCCGTCAGAGCGCGGACATCTTGCCTGTAAAGCC 7730  
OY TTAACCTAAGACTCCGCGCGCGGGCTGCGCTGTGACGACTTAACCTAGGGAGATGTTAAC 7798  
Db 7731 TTAACCTAAGACTCCGCGCGCGGGCTGCGCTGTGACGACTTAACCTAGGGAGATGTTAAC 7790  
OY TGGTCTCGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGG 7858  
Db 7791 TGGTCTCGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGG 7850

OY 7859 ACACACAGGCGGCGACAGGCGGCGACAGGAGCCCAAGACAGATGACCAAGCACTCCAGCGC 7918  
Db 7861 ACACACAGGCGGCGGCGGCGGCGGCGGAGGAGCCCAAGACAGATGACCAAGCACTCCAGCGC 7910  
OY ACTGCTCTCCCGGAATGCAATTTGGAACCAAGTCTTAACTGAGCTGCGACGCTCCGCGC 7978  
Db 7911 ACTGCTCTCCCGGAATGCAATTTGGAACCAAGTCTTAACTGAGCTGCGACGCTCCGCGC 7970  
OY CTCCCTCGGCGCTCCATCCGCTTAGCGCTGACAGATGGAAGCAGAGCCCTGTCCAGC 8038  
Db 7971 CTCCCTCGGCGCTCCATCCGCTTAGCGCTGAGCAATGGAAGCAGAGCCCTGTCCAGC 8030  
OY CCCAGTGCCTGCTTCCGCTTCCCAAGACTGCCCCAGCAACAGATGCTGGAAC 8098  
Db 8031 CCCAGTGCCTGCTTCCGCTTCCCAAGACTGCCCCAGCAACAGATGCTGGAAC 8090  
OY AAGTCAGGCGAGGTGGCGGACAAAGGCGCAGGTGCGGCTGGGGGGAACGATGCTCC 8158  
Db 8091 AAGTCAGGCGAGGTGGCGGACAAAGGCGCAGGTGCGGCTGGGGGGAACGATGCTCC 8150  
OY GAGGACTGGAATGTTTTTTCACATGCTGCGGCGGAGGGGGAAGGAGAGATG 8218  
Db 8151 GAGGACTGGAATGTTTTTTCACATGCTGCGGCGGAGGGGGAAGGAGAGATG 8210  
OY TAAATGATGTGTGTTTACAGGGTATATTTTGAATACCTTCAATGAATTAATTCAGATG 8278  
Db 8211 TAAATGATGTGTGTTTACAGGGTATATTTTGAATACCTTCAATGAATTAATTCAGATG 8270  
OY TTTTACGAGGAAGGAGACTTACCCAGATATTAATGCTGCTGCTTTTGAATCTCTTAC 8338  
Db 8271 TTTTACGAGGAAGGAGACTTACCCAGATATTAATGCTGCTGCTTTTGAATCTCTTAC 8330  
OY GGTTCAGAGGCGGTGAGAGGCGGACAGTGGGTGACCCCATCATCTGCGAGGACCAAGGGG 8398  
Db 8331 GGTTCAGAGGCGGTGAGAGGCGGACAGTGGGTGACCCCATCATCTGCGAGGACCAAGGGG 8390  
OY GCGGGGACTGCTCG-TCAAGCGCCGCTGTGTCTTCCCTCCCTCCCTTCCCTTGGGAGAT 8457  
Db 8391 GCGGGGACTGCTCGTCAAGCGCCGCTGTGTCTTCCCTCCCTCCCTTCCCTTGGGAGAT 8450  
OY GAATTCGATGCGTATTCGTGTGCGCCGCAATTTGCGCAGGGGTGTGTATTCGTCAATTAAC 8517  
Db 8451 GAATTCGATGCGTATTCGTGTGCGCCGCAATTTGCGCAGGGGTGTGTATTCGTCAATTAAC 8510  
OY ACAGTGTCTTAATTAAGCGAATTAATCTCCAAAAAAGGAGGAGGAGGAGGAGGAGGAGG 8571  
Db 8511 ACAGTGTCTTAATTAAGCGAATTAATCTCCAAAAAAGGAGGAGGAGGAGGAGGAGGAGG 8564

RESULT 4  
ACN44283  
ID ACN44283 standard; cDNA, 8667 BP.  
XX ACN44283;  
XX  
AC ACN44283;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human mRNA sequence hCT16428.  
XX  
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN W02003073826-A2.  
XX  
XX 12-SEP-2003.  
PD  
XX 28-FEB-2003; 2003WO-US006235.  
PF  
XX 01-MAR-2002; 2002US-00087192.  
PR  
XX  
XX (SAGR-) SAGRES DISCOVERY.  
PA

XX Morris DW;  
XX WPI; 2003-328604/31.  
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
XX comprises a nucleotide sequence.  
XX  
XX Claim 1; SEQ ID NO 653; Opp; English.  
XX  
XX The present invention relates to novel DNA and protein sequences which  
XX are associated with carcinomas. The sequences are useful for: (i) for  
XX screening drug candidates; (ii) for screening of bioactive agent capable  
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of  
XX a bioactive agent capable of modulating the activity of CAP; (iv) for  
XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
XX carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
XX determining Carcinoma Associated (CA) gene copy number. In addition, the  
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of  
XX carcinoma including lymphoma. The present sequence is one such CA gene  
XX sequence. Note: This patent is an equivalent to basic patent  
XX US2002182586A1, for which no sequence data was published  
SQ  
Sequence 8667 BP; 1860 A; 3066 C; 2575 G; 1166 T; 0 U; 0 Other;  
Query Match 98.8%; Score 8459.2; DB 11; Length 8667;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 8512; Conservative 1; Mismatches 38; Indels 11; Gaps 2;  
QY 1 CATGTCGGGCTCCACACAGCTTGTGACAGACGTGAGGGCCACGTGAGCCCGCTAACCC 60  
DB 115 CATGTCGGGATCCACACAGCTGTGTGCAOAGAGTGAAGGGCCACTGAGCCCGCTAACCC 174  
QY 61 GCCCACAAGCTTTCTTACCCAGTGCAGATCGCCGAGCACAAGAGTGGGCTCTCT 120  
DB 175 GCCCACAAGCTTTCTTACCCAGTGCAGATCGCCGAGCACAAGAGTGGGCTCTCT 234  
QY 121 GAGATGACGACACCACTCCCGGAGTATGCTCCCACTGTGCGGGCTCCATCATCA 180  
DB 225 GAGATGACGACACCACTCCCGGAGTATGCTCCCACTGTGCGGGCTCCATCATCA 294  
QY 181 GCCCAGGCGGAGAGGCTCTCTGCTGTGATGTTCCAGCCCGGGAATGAAGGCTCCA 240  
DB 295 GCCCAGGCGGAGAGGCTCTCTGCTGTGATGTTCCAGCCCGGGAATGAAGGCTCCA 354  
QY 241 GAGAGTCCACTGCGGCGAGAGTCCCACTCATACCTGCGGAGTGGGAAATCAGAGAT 300  
DB 355 GAGAGTCCACTGCGGCGAGAGTCCCACTCATACCTGCGGAGTGGGAAATCAGAGAT 414  
QY 301 GAGATTGATTTAAAGCAAGCGCTCTGAGTGAAGCTGTGAGCTGAGCTGCTGTGAGCC 360  
DB 415 GAGATTGATTTAAAGCAAGCGCTCTGAGTGAAGCTGTGAGCTGAGCTGCTGTGAGCC 474  
QY 361 GTACACCCCTGTGCGCCACGAGGCTGCGGAGTGTGAAGACTTCACAAAGAACCTGAG 420  
DB 475 GTACACCCCTGTGCGCCACGAGGCTGCGGAGTGTGAAGACTTCACAAAGAACCTGAG 534  
QY 421 CCTGACGGGCAAGCTGGAACCGGTGTCTCCCGCAGCGCCCGCACAATGACCTTGAAGT 480  
DB 535 CCTGACGGGCAAGCTGGAACCGGTGTCTCCCGCAGCGCCCGCACAATGACCTTGAAGT 594  
QY 481 GAGAGTGTGCGCGCAAGCTGTCTCCAGAGAGAGTGTATCCAGAACATGAGACCGCTGGA 540  
DB 595 GAGAGTGTGCGCGCAAGCTGTCTCCAGAGAGAGTGTATCCAGAACATGAGACCGCTGGA 654  
QY 541 CGAGAGATCATCATGTGTGAGAGAGATCTTAAGCTGAAGAAAGACAGACAGT 600  
DB 655 CGAGAGATCATCATGTGTGAGAGAGATCTTAAGCTGAAGAAAGACAGACAGT 714  
QY 601 GGAGGAGAGGCTGCCAAGCGCCGAGGCTGAGAGAGCCGTGTCAACCGCCGCTATCGA 660

DB 715 GGAGGAGAGGCTGCCAAGCGCCGAGGCTGAGAAAGCCGTGTCAACCGCCGCTATCGA 774  
QY 661 GTGGAAGACCCGAGCTGTGTGAGATCATCTTACGACGAGAACCGGAAAGAGGCTGAAGC 720  
DB 775 GTGGAAGACCCGAGCTGTGTGAGATCATCTTACGACGAGAACCGGAAAGAGGCTGAAGC 834  
QY 721 TGCACATCGGATTTCTGGAAGGCTGTGGGGCCCAAGTGTGAGAGTCCGCTGTACAAACAGCC 780  
DB 835 TGCACATCGGATTTCTGGAAGGCTGTGGGGCCCAAGTGTGAGAGTCCGCTGTACAAACAGCC 894  
QY 781 CTCGACACCCGCGAGTATCATGAGAACATCAAAATTAACAGGCGATCGGAAAGAGT 840  
DB 895 CTCGACACCCGCGAGTATCATGAGAACATCAAAATTAACAGGCGATCGGAAAGAGT 954  
QY 841 AATCTGTATCTTCAAGAGAGGAATCAAGCTGTGAAACATGAGAACAGAAAGTCTGGCA 900  
DB 955 AATCTGTATCTTCAAGAGAGGAATCAAGCTGTGAAACATGAGAACAGAAAGTCTGGCA 1014  
QY 901 GCGCTATGACAGCTCATGAGGCTTTGAAAAAAGTGGAGCGCATCGAAAAACAACC 960  
DB 1015 GCGCTATGACAGCTCATGAGGCTTTGAAAAAAGTGGAGCGCATCGAAAAACAACC 1074  
QY 961 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGAGATCTACGAAAGCAAGTCCCTGAGAT 1020  
DB 1075 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGAGATCTACGAAAGCAAGTCCCTGAGAT 1134  
QY 1021 CGGCAAGGACCGCGAGCTTCAGAGGCGCATGCAAGACAGAGTGGCGACCGGGCAGTGG 1080  
DB 1135 CGGCAAGGACCGCGAGCTTCAGAGGCGCATGCAAGACAGAGTGGCGACCGGGCAGTGG 1194  
QY 1081 GCTGTTCATGTGTGGCGCGCCCGAGCAGACAGAGGTGTGAGATCATGTGAGCTCTTC 1140  
DB 1195 GCTGTTCATGTGTGGCGCGCCCGAGCAGACAGAGGTGTGAGATCATGTGAGCTCTTC 1254  
QY 1141 AAGACGAGAGAACTTGAGAGACAGATGCGCAAGTGGCTGTATCCGCCCATGCTGTGA 1200  
DB 1255 AAGACGAGAGAACTTGAGAGACAGATGCGCAAGTGGCTGTATCCGCCCATGCTGTGA 1314  
QY 1201 CGAGCTGACACGACAGCGCATCAAGTTTCATCAACATGAAACGGGCTTTATGGCGGACCCCAT 1260  
DB 1315 CGAGCTGACACGACAGCGCATCAAGTTTCATCAACATGAAACGGGCTTTATGGCGGACCCCAT 1374  
QY 1261 GAAGGTGTCAAAAGACCGCCAGGTCATGAACATGTGAGTGAAGCAGAGAAAGAGACTTT 1320  
DB 1375 GAAGGTGTCAAAAGACCGCCAGGTCATGAACATGTGAGTGAAGCAGAGAAAGAGACTTT 1434  
QY 1321 CGGGAAGAAAGTTTATGACAGCATCCCAAGAACTTTGGCTGATTCGATCATTTCTGTGAGAG 1380  
DB 1435 CGGGAAGAAAGTTTATGACAGCATCCCAAGAACTTTGGCTGATTCGATCATTTCTGTGAGAG 1494  
QY 1381 GAAGACATGTGTGAGTGTGCTCTCTATTATCTTACTGACTTAAGAAATGAAGAACTTTAA 1440  
DB 1495 GAAGACATGTGTGAGTGTGCTCTCTATTATCTTACTGACTTAAGAAATGAAGAACTTTAA 1554  
QY 1441 GAGCTGTGTAAGCGGAGCTTATCGGCGCGCGCAAGAGCCAGCAGCAACAAGCAGCA 1500  
DB 1555 GAGCTGTGTAAGCGGAGCTTATCGGCGCGCGCAAGAGCCAGCAGCAACAAGCAGCA 1614  
QY 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
DB 1615 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1674  
QY 1561 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
DB 1675 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1734  
QY 1621 CGACAAAGAAAGCTCTCTCAAGAGAAAGCAGCAGCACTCAGGAGAGAGAGAGAGAGAG 1680  
DB 1735 CGACAAAGAAAGCTCTCTCAAGAGAAAGCAGCAGCACTCAGGAGAGAGAGAGAGAGAG 1794  
QY 1681 GAAGAGGCTGTGTGCTCTCAAAAGGCTCGCAAACTGTCCAAACAGCCAGAGAGAGAGAG 1740  
DB 1795 GAAGAGGCTGTGTGCTCTCAAAAGGCTCGCAAACTGTCCAAACAGCCAGAGAGAGAGAG 1854

QY 1741 CCGCATCACCCGCTCAATGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCAGCA 1800  
DB 1855 CCGCATCAACCGGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCAGCA 1914  
QY 1801 GAGCCCGGAGCTGGCCCTCCATGGAAGCTGATGAGAGTTCTCCCTGAGCAGAAAGAAAT 1860  
DB 1915 GAGCCCGGAGCTGGCCCTCCATGGAAGCTGATGAGAGTTCTCCCTGAGCAGAAAGAAAT 1974  
QY 1851 GGAACAGCGAAGAAAGGCTCTCTGAAACAAGCGCGCAACTGTGGCCATGCGCCGAT 1920  
DB 1975 GGAACAGCGAAGAAAGGCTCTCTGAAACAAGCGCGCAACTGTGGCCATGCGCCGAT 2034  
QY 1921 GGTGGGCTTCAAGACTGTGTCCAGTGTAAAGACTTCACTTCAACTAACAGAAAGGCA 1980  
DB 2035 GGTGGGCTTCAAGACTGTGTCCAGTGTAAAGACTTCACTTCAACTAACAGAAAGGCA 2094  
QY 1981 GAACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAAGATGAGAAAGAGAGAGAGCGCGC 2040  
DB 2095 GAACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAAGATGAGAAAGAGAGAGAGCGCGC 2154  
QY 2041 GAGGAAGAAAGAAAGAGCGCGCGCGCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
DB 2155 GAGGAAGAAAGAAAGAGCGCGCGCGCGCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2214  
QY 2101 GAGAGATGAGAGATGAGAGCGCTCGGCGCTGAGCGGAAATGAGAGAGAGATGAGAGAG 2160  
DB 2215 GAGAGATGAGAGATGAGAGCGCTCGGCGCTGAGCGGAAATGAGAGAGAGATGAGAGAG 2274  
QY 2161 GGCTGAAGCTTTAATGCTCTCTGAGAAATGAGGTGCCAGAGGGGAAATGAGATGGCCCAAC 2220  
DB 2275 GGCTGAAGCTTTAATGCTCTCTGAGAAATGAGGTGCCAGAGGGGAAATGAGATGGCCCAAC 2334  
QY 2221 CACTGTCAACAACACTGACAGACACCGAGAGCATGCCCTCTCTCTCACTCTGAGCGCGCAA 2280  
DB 2335 CACTGTCAACAACACTGACAGACACCGAGAGCATGCCCTCTCTCTCACTCTGAGCGCGCAA 2394  
QY 2281 GGAACAAGGGCGAAATGGGCGCCAGCGCCCAAGCCCTGGGCGCGCAAGCGCGCAACCCCG 2340  
DB 2395 GGAACAAGGGCGAAATGGGCGCCAGCGCCCAAGCCCTGGGCGCGCAAGCGCGCAACCCCG 2454  
QY 2341 AGGCCCAACCAACCCCAACCGAGAGACATCCGGGGCCCCCATTTAGGCCCAACCCGGGCTTC 2400  
DB 2455 AGGCCCAACCAACCCCAACCGAGAGACATCCGGGGCCCCCATTTAGGCCCAACCCGGGCTTC 2514  
QY 2401 TGAAGCAACCGGAGCCCTTACGCCCTCAACAGACCCCATGCGCTTGTGACCTCTCC 2460  
DB 2515 TGAAGCAACCGGAGCCCTTACGCCCTCAACAGACCCCATGCGCTTGTGACCTCTCC 2574  
QY 2461 TGTGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
DB 2575 TGTGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2634  
QY 2521 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
DB 2635 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2694  
QY 2581 GCCCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
DB 2695 GCCCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2754  
QY 2641 GAGAGCCGCTGAGGCAACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
DB 2755 GAGAGCCGCTGAGGCAACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2814  
QY 2701 GGGCAGGGCCACCACTGCGCAAGAGCTCGGGCGCCCCCAGAGACAGGACTTCAAGTGTAC 2760  
DB 2815 GGGCAGGGCCACCACTGCGCAAGAGCTCGGGCGCCCCCAGAGACAGGACTTCAAGTGTAC 2874  
QY 2761 CTGCGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
DB 2875 CTGCGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2934

QY 2821 AAGGCCAGCCTCTCTCAACCCGAGCTGGCGACCCCGGGCCAAATGCTTCAACCCAGAGAGCC 2880  
DB 2935 AAGGCCAGCCTCTCTCAACCCGAGCTGGCGACCCCGGGCCAAATGCTTCAACCCAGAGAGCC 2994  
QY 2881 ACTGGAACCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
DB 2995 ACTGGAACCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3054  
QY 2941 AGTCATGAGCCCGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000  
DB 3055 AGTCATGAGCCCGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3114  
QY 3001 GCCACCGCAAAACCTTGACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
DB 3115 GCCACCGCAAAACCTTGACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3174  
QY 3061 GGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120  
DB 3175 GGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3234  
QY 3121 GAAGCTGCTGGGGAGACCCCGCTTGAGACTTCCGGCTGACCTTCCCGTGCACCCCGCG 3180  
DB 3235 GAAGCTGCTGGGGAGACCCCGCTTGAGACTTCCGGCTGACCTTCCCGTGCACCCCGCG 3294  
QY 3181 TGAGGTGATCAAGGCTTCCCGCATGCCCCGAGACCTCTGAGCTTCTCTTACGCTTCAAC 3240  
DB 3295 TGAGGTGATCAAGGCTTCCCGCATGCCCCGAGACCTCTGAGCTTCTCTTACGCTTCAAC 3354  
QY 3241 TGTGTACCACTGCGCCCGGGGCTTCATGACATGCGCGGCGCGCTGCGCGCGCGCGCAC 3300  
DB 3355 TGTGTACCACTGCGCCCGGGGCTTCATGACATGCGCGGCGCGCTGCGCGCGCGCGCAC 3414  
QY 3301 CACCATCTCAACCCCGCTCCCTCATCTCTCTGCGCAAGACCCCAAGAGAGAGAGAGAGAG 3360  
DB 3415 CACCATCTCAACCCCGCTCCCTCATCTCTCTGCGCAAGACCCCAAGAGAGAGAGAGAGAG 3474  
QY 3361 GCAAAATAGGTGCATCTTCCCAAGAAATGCGGTCCAGCTTCCAGCTTCCGATCTCAGAGCA 3420  
DB 3475 GCAAAATAGGTGCATCTTCCCAAGAAATGCGGTCCAGCTTCCAGCTTCCGATCTCAGAGCA 3534  
QY 3421 TGCCAAAGGCGCCCGGGGCGCTGTCAACATGAGGAGTGCGCCGAGCCATGAGAGAGAGAGAG 3480  
DB 3535 TGCCAAAGGCGCCCGGGGCGCTGTCAACATGAGGAGTGCGCCGAGCCATGAGAGAGAGAGAG 3594  
QY 3481 GCTGCGCAACCTTTCAGCGAGTGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540  
DB 3595 GCTGCGCAACCTTTCAGCGAGTGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3654  
QY 3541 ACCGAGAGAGCTGGGGGTGCCCAAGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600  
DB 3655 ACCGAGAGAGCTGGGGGTGCCCAAGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3714  
QY 3601 GGGCTCAATTCGCGCGGAGAGAGATCAACCAAGAGCATTCCTCCAGCAACGAGGTGCCCTCGA 3660  
DB 3715 GGGCTCAATTCGCGCGGAGAGATCAACCAAGAGCATTCCTCCAGCAACGAGGTGCCCTCGA 3774  
QY 3661 CAGCGCATCAATTCAGCGGCTTCATCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720  
DB 3775 CAGCGCATCAATTCAGCGGCTTCATCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3834  
QY 3721 GGGCAGCATCAACAGAGATCATTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780  
DB 3835 GGGCAGCATCAACAGAGATCATTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3894  
QY 3781 GGAAGAGCTTGGCCCAAGGGGCAAGTCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3840  
DB 3895 GGAAGAGCTTGGCCCAAGGGGCAAGTCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3954  
QY 3841 TGAAGGTGAGCATGTCTGTGACCAAGTGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3900  
DB 3955 TGAAGGTGAGCATGTCTGTGACCAAGTGTCTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAG 4014  
QY 3901 CCCCCATGAGAGAGCGCGCGCCCCCAAGCGGACCTTATGACATGATGAGAGAGCGCGGTGGCAG 3960

Db	4015	CCCCATBAGACGGCGCCCCCGACCCCTATGACATGATGAGGGCCCGCTGGGACG	4074
Qy	3961	AGCCATCTCTCAGCCAGCATCGAAGGTCTCATYGGGCGTGCACATCCCGCGGACGACA	4020
Db	4075	AGCCATCTCTCAGCCAGCATCGAAGGTCTCATYGGGCGTGCATCCCGCGGACGACA	4134
Qy	4021	CAGCCCCCACCACCTCAAAAGACAGCACACATCCGGGGTTCATACACAAAGGATCC	4080
Db	4135	CAGCCCCCACCACCTCAAAAGACAGCACACATCCGGGGTTCATACACAAAGGATCC	4194
Qy	4081	TCGGTCTCACTGAGAGGACAGAGGAATTACTTCGTTCGGGAGGCCAAAGTCTTAAAGCG	4140
Db	4195	TCGGTCTCACTGAGAGGACAGAGGAATTACTTCGTTCGGGAGGCCAAAGTCTTAAAGCG	4254
Qy	4141	GGAAGGACAGCGCTCCGCCCCCACCAGCGCTTCACGGGACTTGACCGAGGCTTCAAGAAGCA	4200
Db	4255	GGAAGGACAGCGCTCCGCCCCCACCAGCGCTTCACGGGACTTGACCGAGGCTTCAAGAAGCA	4314
Qy	4201	GGCCCTGGGGCCCCCTGAAAGCTGAAGCCGGGCCCATGAGGGGCTGGTGGCCACGGTGAAGA	4260
Db	4315	GGCCCTGGGGCCCCCTGAAAGCTGAAGCCGGGCCCATGAGGGGCTGGTGGCCACGGTGAAGA	4374
Qy	4261	GGCGGGCGCTCTCATTCATYGAATTCGGCGGGAAGAGCTGGGGACACAGCCCGAGCTGCC	4320
Db	4375	GGCGGGCGCTCTCATTCATYGAATTCGGCGGGAAGAGCTGGGGACACAGCCCGAGCTGCC	4434
Qy	4321	CCTGGAGCCCGGGCGGCTCAAGAGAGGGTCCATACAGCAGGGGACCCCGCTCAAGTACGA	4380
Db	4435	CCTGGAGCCCGGGCGGCTCAAGAGAGGGTCCATACAGCAGGGGACCCCGCTCAAGTACGA	4494
Qy	4381	CACCGGCGCGTTCACACACTGGCTCCAAAGAACAGACGTAAGTCTCTCATCCGACAGCC	4440
Db	4495	CACCGGCGCGTTCACACACTGGCTCCAAAGAACAGACGTAAGTCTCTCATCCGACAGCC	4554
Qy	4441	CGGCGGACGTTCCCAACCCGTGCACCCGCTGGAATGTGATGSCCGACCGCCGGGCACTGGA	4500
Db	4555	CGGCGGACGTTTCCCAACCCGTGCACCCGCTGGAATGTGATGSCCGACCGCCGGGCACTGGA	4614
Qy	4501	ACGNGCCCTGCATACAGAGAGAGCCCTGAAGAGCCGGCCAGGGACCCGACAGACTGGGGGG	4560
Db	4615	ACGNGCCCTGCATACAGAGAGAGCCCTGAAGAGCCGGCCAGGGACCCGACAGACTGGGGGG	4674
Qy	4561	CTCCATTTGCGGCGGCGCCCGCGTCAATYGTGCTCTGAAGCTGGTTAAGCGCGCGGACAGCC	4620
Db	4675	CTCCATTTGCGGCGGCGCCCGCGTCAATYGTGCTCTGAAGCTGGGGCAAGCCGCGGACAGCC	4734
Qy	4621	CTGACCTATGAGGACCAACGGGGGCAACCTTTTGGCGGCCACTTCCACAGAGGTTGCGCGT	4680
Db	4735	CTTAACTATGAGGACCAACGGGGGCAACCTTTTGGCGGCCACTTCCACAGAGGTTGCGCGT	4794
Qy	4681	GACCATGCGGGAGGCCACGCGCGGCTGACAGAGAGGAGCCTTTTCCTCAGCAAGGATC	4740
Db	4795	GACCATGCGGGAGGCCACGCGCGGCTGACAGAGAGGAGCCTTTTCCTCAGCAAGGATC	4854
Qy	4741	CCAGGACCGAAGCTGAGCTGACGCGCTCGTGAATGSCCAAAGTCCCGGACAGACACGCT	4800
Db	4855	CCAGGACCGAAGCTGAGCTGAGCGCTCGTGAATGSCCAAAGTCCCGGACAGACACGCT	4914
Qy	4801	GCCGAGACCAACCAACCCCATCTGCGCCCTATGAGCACTGTGCTTGGGGGCTGAATGG	4860
Db	4915	GCCGAGACCAACCAACCCCATCTGCGCCCTATGAGCACTGTGCTTGGGGGCTGAATGG	4974
Qy	4861	CGTGGACCTTGATGCGACCCACATTCCTCCCTTGGCTTGCACCCCACTTCATATCCCGCGG	4920
Db	4975	CGTGGACCTTGATGCGACCCACATTCCTCCCTTGGCTTGCACCCCACTTCATATCCCGCGG	5034
Qy	4921	CATCCCTCTGAGCAGAGCGCTGCCTACTTACTGCGCCGCAACCTGAGGCCCAACCCGAC	4980
Db	5035	CATCCCTCTGAGCAGAGCGCTGCCTACTTACTGCGCCGCAACCTGAGGCCCAACCCGAC	5094
Qy	4981	CTACCCGCACTGTACCCACCTTACTTATCGCGGCTTACCCGCAACGCGGCGCTTGA	5040

Db	5095	CTACCGCAGCACTGTATACCAACCCCTACCTCTATCCGGGGCTTACCCCGACACAGGGGGCGCTGGA	5154
Qy	5041	GAACCGGCAGACCATCATCATATGACTACATCACTCCGACGATGTCACCAACACCGC	5100
Db	5155	GAACCGGCAGACCATCATCATGACTACATCACTCCGACGATGTCACCAACCGCGGC	5214
Qy	5101	CACCGGCATATGGCCCAAGGAAGTGAATATGTGAAGGGGGCTCTCGGCCCGCGAGTCTCGCT	5160
Db	5215	CACCGGCATATGGCCCAAGGAAGTGAATATGTGAAGGGGGCTCTCGGCCCGCGAGTCTCGCT	5274
Qy	5161	GGCACTCAACTACGCTGCGGGGTCCCGAGGCAATCATGACCTGTGCCAATGGCACACCT	5230
Db	5275	GGCACTCAACTACGCTGCGGGGTCCCGAGGCAATCATGACCTGTGTGCCAATGGCACACT	5334
Qy	5221	GCCTGTGCTGTGCTCCCGGACACCAAGGACCCCGACCGCAATGGAACCGCTTGTCTTA	5280
Db	5335	GCCTGTGCTGTGCTCCCGGACACCAAGGACCCCGACCGCAATGGAACCGCTTGTGCTTA	5394
Qy	5281	CTTCCCAACCGCGCGCCCGACGCTTACAGAGCCGTCACAGACGCTTCCCACTTCTCCAGG	5340
Db	5395	CTTCCCAACCGCGCGCCCGACGCTTACAGAGCCGTCACAGACGCTTCCCACTTCTCCAGG	5454
Qy	5341	AGGTCCAAACAACATTGACAAACCAACCAACAGCTCTCGTCCGAGGGGGAGGAGACCG	5400
Db	5455	AGGTCCAAACAACATTGACAAACCAACCAACAGCTCTCGTCCGAGGGGGAGGAGACCG	5514
Qy	5401	GGATCGAGAGCGGAGACCGGATCGGAGCGGGAATAATCTCATCTCTCACTCCACACGAC	5460
Db	5515	GGATCGAGAGCGGAGACCGGATCGGAGCGGGAATAATCTCATCTCTCACTCCACACGAC	5574
Qy	5461	GGTGGAGCAGCAGCACCCATCTGGAGACCTGGTACAGAGCAGACGCGCAGCAGCGCGAG	5520
Db	5575	GGTGGAGCAGCAGCACCCATCTGGAGACCTGGTACAGAGCAGACGCGCAGCAGCGCGAG	5634
Qy	5521	CAGCGGGGGGGGTGGGGGACAGCAGACCGCCCGCCCTCCCACTCCCATGCGCCACAGCA	5580
Db	5635	CAGCGGGGGGGGTGGGGGACAGCAGACCGCCCGCCCTCCCACTCCCATGCGCCACAGCA	5694
Qy	5581	CTCGCCCATCTCCCTCTGGACCCGAGATGCGCTTCAAGACAGACCCAGTGTCTTCAAA	5640
Db	5695	CTCGCCCATCTCCCTCTGGACCCGAGATGCGCTTCAAGACAGACCCAGTGTGTTCACAA	5754
Qy	5641	CACAGGATGAAGGGTATCATCAACCGTGTGAGACCCAGCAAGCCACGCTCTGAGGTC	5700
Db	5755	CACAGGATGAAGGGTATCATCAACCGTGTGAGACCCAGCAAGCCACGCTCTGAGGTC	5814
Qy	5701	CACCTCCACTCTCTCAACCGGTTGGCCCAAGTGTGCACATTTCCCACTGTGCACACCACTGCC	5760
Db	5815	CACCTCCACTCTCTCAACCGGTTGGCCCGGCTGTGCACATTTCCCACTGTGCACACCACTGCC	5874
Qy	5761	ACTGGGGGGGACCTCTGATGGGGGTCTAACCTTACCTCATGGAAGCCGCTCTTGTGCGCCAA	5820
Db	5875	ACTGGGGGGGACCTCTGATGGGGGTCTAACCTTACCTCATGGAAGCCGCTCTTGTGCGCCAA	5934
Qy	5821	GGAGGCCCTCCCGGGTGTGCCCGGACAGAGCGGACCCCGAGCAGACACCGGCAATGTGCTTCT	5880
Db	5935	GGAGGCCCTCCCGGGTGTGCCCGGACAGAGCGGACCCCGAGCAGACACCGGCAATGTGCTTCT	5994
Qy	5881	CGCCAGCCCCCAGCCCCGCTCTCGGGCTGAGACCCGCTCTCTCCCCAGCAAGGGCTGTGGA	5940
Db	5995	CGCCAGCCCCCAGCCCCGCTCTCGGGCTGAGACCCGCTCTCTCCCCAGCAAGGGCTGTGGA	6054
Qy	5941	GGCCCCGGCCCTCTATGTGCTCTGTCTCTGTGGACAGCACACATGTGCGCGCACCCCTGTGGAA	6000
Db	6055	GGCCCCGGCCCTCTATGTGCTCTGTCTCTGTGGACAGCACACATGTGCGCGCACCCCTGTGGAA	6114
Qy	6001	GAACCTTGCACTTCAACACGCGCAGGCCCGGACCCGCGGCGCACCTTGCTCGGCTTGTGA	6060
Db	6115	GAACCTTGCACTTCAACACGCGCGAGCCCGGACCCGCGGCGCACCTTGCTCGGCTTGTGA	6174
Qy	6061	CCCGCACCGGGAAAAAGACTAAAGTAAACCTTTTTCATCTCAGGAACCTGGAATCTCGCTTC	6120
Db	6175	CCCGCACCGGGAAAAAGACTAAAGTAAACCTTTTTCATCTCAGGAACCTGGAATCTCGCTTC	6234



QY 6121 TCTGGGTTAACACGCGACAGCTACAGCCCGAAGGGGTGAGCCCGTCAAGCCCTGTGAG 6180  
DB 6235 TCTGGGTTAACACGCGACAGCTACAGCCCGAAGGGGTGAGCCCGTCAAGCCCTGTGAG 6294  
QY 6181 CTCAACCGAGTCTGACACCGACGACGAAGGGGTCTCCCAAGGACCTTGGAAGAGCTGACAAAGG 6240  
DB 6295 CTCAACCGAGTCTGACACCGACGACGAAGGGGTCTCCCAAGGACCTTGGAAGAGCTGACAAAGG 6354  
QY 6241 CCACCTGAGGGGGAGCTGCGGCGCCAAAGCAGCAGGGCCCGTGAAGCTTGGCGGGAGGCG 6300  
DB 6355 CCACTTGAGGGGGAGCTGCGGCGCCAAAGCAGCAGGGCCCGTGAAGCTTGGCGGGAGGCG 6414  
QY 6301 CGCCCACTCCCAACCTGCGGCGCGTGTGAGAGCCAGCCCTGTGTCAAGCCCGTGTCT 6360  
DB 6415 CGCCCACTCCCAACCTGCGGCGCGTGTGAGAGCCAGCCCTGTGTCAAGCCCGTGTCT 6474  
QY 6381 CCAAGACCGGCGGAGGGGTCAAGGTCAACAGGGGGGTGTCAACCTTGAGCCCAACATTCAG 6420  
DB 6475 CCAAGACCGGCGGAGGGGTCAAGGTCAACAGGGGGGTGTCAACCTTGAGCCCAACATTCAG 6534  
QY 6421 TGAAGTCAATCAACACAGACTACACCGGACACACCAACAGACTCAGGCGACCCCTGCG 6480  
DB 6535 TGAAGTCAATCAACACAGACTACACCGGACACACCAACAGACTCAGGCGACCCCTGCG 6594  
QY 6481 CGCCCCCTCTATCTCCCTTCCCTGGGGCGACGCTGCCCGTCTTGAGACTTCGCGCGCCACG 6540  
DB 6595 CGCCCCCTCTATCTCCCTTCCCTGGGGCGACGCTGCCCGTCTTGAGACTTCGCGCGCCACG 6654  
QY 6541 CAGTACCTCTATCTCCCGGCGGCGGACCATGTGTCCCGGCGGCGGTCTCCCGCCACAG 6600  
DB 6655 CAGTACCTCTATCTCCCGGCGGCGGACCATGTGTCCCGGCGGCGGTCTCCCGCCACAG 6714  
QY 6601 CGAAGGGGGCAAGAGGTCTCCAGAGCAAAAGAGTGTGAGTGTGGGTGTGAGGGA 6660  
DB 6715 CGAAGGGGGCAAGAGGTCTCCAGAGCAAAAGAGTGTGAGTGTGGGTGTGAGGGA 6774  
QY 6661 CGGTATTGAACCTGTGTCCCAACCGAGGGGATGACGAGCCAGGGCACTCCCGAGTGC 6720  
DB 6775 CGGTATTGAACCTGTGTCCCAACCGAGGGGATGACGAGCCAGGGCACTCCCGAGTGC 6834  
QY 6721 TGTGTACCCCGCTGTGTACCCGGGATGGGGAAACAGACGAGCCCAAGATGGGCTCCCA 6780  
DB 6835 TGTGTACCCCGCTGTGTACCCGGGATGGGGAAACAGACGAGCCCAAGATGGGCTCCCA 6894  
QY 6781 GTCTCCAGGCAACACAGCCAGCCGCGACGCTTCTTCAAGCAAGCTGACCGAAGCACTC 6840  
DB 6895 GTCTCCAGGCAACACAGCCAGCCGCGACGCTTCTTCAAGCAAGCTGACCGAAGCACTC 6954  
QY 6841 CGCCATGTGCAAGTCCAGAGCAAGAGATCAACAAAGAGCTGAAACCCCAACCCGGA 6900  
DB 6955 CGCCATGTGCAAGTCCAGAGCAAGAGATCAACAAAGAGCTGAAACCCCAACCCGGA 7014  
QY 6901 TGAAGCTGAATCAATATTCAGGCAAGCTTGGGAGCGAGATCTTCAATATGCGCCATCAG 6960  
DB 7015 TGAAGCTGAATCAATATTCAGGCAAGCTTGGGAGCGAGATCTTCAATATGCGCCATCAG 7074  
QY 6961 CGGAACAGGCTTATGACCTATAGAGAGCAGAGCGGTGAGAAACATGCGACGACCAACAT 7020  
DB 7075 CGGAACAGGCTTATGACCTATAGAGAGCAGAGCGGTGAGAAACATGCGACGACCAACAT 7134  
QY 7021 GGGGCTGAGGCGCATATTTAGAAAGCACTCATGGGTAAATATGACAGATGGGAAAGTGC 7080  
DB 7135 GGGGCTGAGGCGCATATTTAGAAAGCACTCATGGGTAAATATGACAGATGGGAAAGTGC 7194  
QY 7081 CCGCGCGCTCAGGCGCAATGCTTTTAACTCTGTGAATGCGCATGCGAGCTGCCGCTGC 7140  
DB 7195 CCGCGCGCTCAGGCGCAATGCTTTTAACTCTGTGAATGCGCATGCGAGCTGCCGCTGC 7254  
QY 7141 TATGCCCATTAACCGCTGTGACGAGACGAGTGAACCACTCACTCGCGCAGGTGCGG 7200  
DB 7255 TATGCCCATTAACCGCTGTGACGAGACGAGTGAACCACTCACTCGCGCAGGTGCGG 7314

QY 7201 CGGGAAGGCCAAGGTCTCTGGCAGACCCAGACCCGAAAGCCAAAGTCCCGGCCCCGGG 7260  
DB 7315 CGGGAAGGCCAAGGTCTCTGGCAGACCCAGACCCGAAAGCCAAAGTCCCGGCCCCGGG 7374  
QY 7261 CCTGGCATCTGGGGAGCCGGGCCAACCCTGTCTCCTCAGTGCATGTGAGGGAGACTGCA 7320  
DB 7375 CCTGGCATCTGGGGAGCCGGGCCAACCCTGTCTCCTCAGTGCATGTGAGGGAGACTGCA 7434  
QY 7321 CGGCGGAGCGCGCTCAACCAACCGGTGTGGAGAGACAGGCCCTGTCTCGCAGTTCAC 7380  
DB 7435 CGGCGGAGCGCGCTCAACCAACCGGTGTGGAGAGAGAGGCCCTGTCTCGCAGTTCAC 7494  
QY 7381 GCCATTCCCTTACAAACCCCTGATATGATGCGGTGACGGCGGTGTATGAGTCTCCAC 7440  
DB 7495 GCCATTCCCTTACAAACCCCTGATATGATGCGGTGACGGCGGTGTATGAGTCTCCAC 7554  
QY 7441 CCCACGGGCGTCCCGCGGGAGCGGGCCCTCGCTGAGCCCGCACAGCTTGAGAGCA 7500  
DB 7555 CCCACGGGCGTCCCGCGGGAGCGGGCCCTCGCTGAGCCCGCACAGCTTGAGAGCA 7614  
QY 7501 GAGGCCAAGCCACTGTCTGTGCAAGTACGAGACACTTTCGACAGCGATGACTGAG 7560  
DB 7615 GAGGCCAAGCCACTGTCTGTGCAAGTACGAGACACTTTCGACAGCGATGACTGAG 7674  
QY 7561 AACAGGGCGGGGGGGCGGGCGGTGEBURSPERLUTGACAGGTCCAGCGAGCCACAGAA 7620  
DB 7675 AACAGGGCGGGGGGGGGGGGG-----TGTAGGTTCCAGCGAGCCACAGAA 7724  
QY 7621 CGGCGCTCAGAGAGCGGGGGCGGCTGCGACTCCCGCAACCAAGAAAGAGCCCTGAGTC 7680  
DB 7725 CGGCGCTCAGAGAGCGGGGGCGGCTGCGACTCCCGCAACCAAGAAAGAGCCCTGAGTC 7784  
QY 7681 CGCCTGCGCTTCATCATCTGTCCGTCCAGACCGGAGTCTTGCCTGTCTAAAGCTTT 7740  
DB 7785 CGCCTGCGCTTCATCATCTGTCCGTCCAGACCGGAGTCTTGCCTGTCTAAAGCTTT 7844  
QY 7741 AACTAAGACTCCCGCGCGGGGTGAGCCTGTGACAACTTACTAGGGGATGTTTAACTG 7800  
DB 7845 AACTAAGACTCCCGCGCGGGGTGAGCCTGTGACAACTTACTAGGGGATGTTTAACTG 7904  
QY 7801 GTGCTCGGGAAGGAGGGGGAAGGGGCGGGGAGGGGGGACCGCAGGCGTGTGGAGCCAC 7860  
DB 7905 GTGCTCGGGAAGGAGGGGGAAGGGGCGGGGAGGGGGGACCGCAGGCGTGTGTGAGCCAC 7964  
QY 7861 ACAAGGCGGCGAGGCGGCGGACGAGACCCAAAGCAGATGACACAGCACTTCCAGCCAC 7920  
DB 7965 ACGAGGCGGCGAGGCGGCGGACGAGACCCAAAGCAGATGACACAGCACTTCCAGCCAC 8024  
QY 7921 TGCTTCCCCCGAATGCAATTTTGAACCAAGTCTTAACTGAGTGCAGACCCCGGCGCT 7980  
DB 8025 TGCTTCCCCCGAATGCAATTTTGAACCAAGTCTTAACTGAGTGCAGACCCCGGCGCT 8084  
QY 7981 CCTTCGCGCTCCATCCCGCTTGAAGCTGTGACAGATGAGACGAGGCGCTGTCCAGCCC 8040  
DB 8085 CCTTCGCGCTCCATCCCGCTTGAAGCTGTGACAGATGAGACGAGGCGCTGTCCAGCCC 8144  
QY 8041 CCAAGTGCCTGTTTCCGATCCCAAGACTGCGCCCAAGCAAGAGATTTCTGGAACCA 8100  
DB 8145 CCAAGTGCCTGTTTCCGATCCCAAGACTGCGCCCAAGCAAGAGATTTCTGGAACCA 8204  
QY 8101 GTCAGGCGCAGGTGGCGGACAAAGGGCCAGGTGCGGCTTGGGGGGAAACGATGCTCCGA 8160  
DB 8205 GTCAGGCGCAGGTGGCGGACAAAGGGCCAGGTGCGGCTTGGGGGGAAACGATGCTCCGA 8264  
QY 8161 GGAAGTGAATGTTTTTCAACATCGTTTGCGAGCGGTGGGAAAGGAAGGACAGATGA 8220  
DB 8265 GGAAGTGAATGTTTTTCAACATCGTTTGCGAGCGGTGGGAAAGGAAGGACAGATGA 8324  
QY 8221 AATGATGTGTTGTTTACAGGGTATATTTTGAATCTTCAATGAATTAATTCAATGTT 8280  
DB 8325 AATGATGTGTTGTTTACAGGGTATATTTTGAATCTTCAATGAATTAATTCAATGTT 8384  
QY 8281 TTAAGCAAGGAAGACTTAACCAAGTATTAATGCTGCTGTGCTTTTGAATCTGCTTACCG 8340

[illegible]

## RESULT 5

ID ADQ84524 standard; cDNA; 8533 BP

AC ADQ84524;

DT 07-OCT-2004 (first entry)

DE	Human tumour-associated antigenic target (TAT) cDNA sequence #1338.
25	

human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;

AM CANCER; CELL PROLIFERATIVE UNBOUNDED; GENES; BR.

**Homo sapiens.**

PN WO2004060270-A2.

PD 22-JUL-2004

PF 15-OCT-2003; 2003WO-US029126.

PR 18-OCT-2002; 2002US-0418988P.

PA (GETH ) GENENTECH INC.

PA (ZHOU/) ZHOU Y.

Wu TD, Zhou Y;

DR WPI; 2004-534300/51.

PT New nucleic acid molecule and encoded polypeptide, for diagnosing,

[illegible][illegible]

CC The present invention describes an isolated tumour-associated antigenic  
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide  
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of  
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%  
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-  
CC (c). Also described: (1) an expression vector comprising the above  
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)  
CC a process for producing a polypeptide; (4) an isolated polypeptide  
CC comprising: (a) an amino acid sequence encoded by any of the above  
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-  
CC length coding region of the above nucleotide sequences; or (c) a sequence  
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide  
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)  
CC an isolated antibody that binds to the above polypeptide; (7) a process  
CC for producing the antibody; (8) an isolated oligopeptide that binds to  
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)  
CC binding organic molecule that binds to the above polypeptide; (10) a

composition of matter comprising the above (chimeric) polypeptide, antibody, oligopeptide or TAT binding organic molecule, in combination with a carrier; (11) an article of manufacture comprising a container and the composition of matter contained within the container; (12) methods of inhibiting the growth of a cell that expresses the above protein, where the growth of the cell is at least in part dependent upon a growth potentiating effect of the above protein; (13) a method of therapeutically treating a mammal having a cancerous tumor comprising cells that express the above protein; (14) a method of determining the presence of a protein in a sample suspected of containing the protein described above; (15) methods of diagnosing the presence of a tumor in a mammal; (16) a method for treating or preventing a cell proliferative disorder associated with increased expression or activity of the above protein; and (17) a method of binding an antibody, oligopeptide or organic molecule to a cell that expresses the protein described above. The TAT sequences have cytostatic activities, and can be used in gene therapy. The composition and methods are useful for diagnosing, preventing or treating cancer. The composition is also used for preparing a medicament for the therapeutic treatment or diagnostic detection of a cell proliferative disorder or cancer. The present sequence represents a human TAT cDNA sequence from the present invention.

Seq 8533 BP; 1836 A; 3034 C; 2528 G; 1135 T; 0 U; 0 Other;

Query Match	98.3%	Score 8423.4	DB 13	Length 8533
Best Local Similarity	99.3%	Pred. No. 0		
Matches 8493; Conservative	3	Mismatches	36; Indels	22; Gaps 3

QY	1	CATGTCGGGGCTCCCA	CA	CAGGCTTGGTGG	CAGACCGTGGAGGGG	CACGTAGAGCCCGCATACC	60
Db	1	CATGTGGGAATCCACA	CAGGCTGTGTGG	CACAGCTGTGGAGGGG	CACGTAGAGCCCGCATACC	60	
QY	61	GCCCCA	CAGCCTTTCTTA	CCAGTGGCAGATGG	CCCGGAGCGACA	CGAGCGTGGGGTCTCT	120
Db	61	GCCCCA	CAGCCTTTCTTA	CCAGTGGCAGATGG	CCCGGAGCGACA	CGAGCGTGGGGTCTCT	120
QY	121	GGAAATCA	CAGACA	CAACTTCCCGGAGTAA	TGCTTCCACCTGTGCGCGGGTCC	ATATCCAA	180
Db	121	GGAAATCA	CAGACA	CAACTTCCCGGAGTAA	TGCTTCCACCTGTGCGCGGGTCC	ATATCCAA	180
QY	181	GCCCCA	CGGGGAGGAGCC	CTCCCTGCTGTGAGT	TTCAGCCCGGGAAATGA	ACGATCCCA	240
Db	181	GCCCCA	CGGGGAGGAGCC	CTCCCTGCTGTGAGT	TTCAGCCCGGGAAATGA	ACGATCCCA	240
QY	241	GGAGCTCA	CCTGCGGCCA	GAGTCCCACTA	TACTCTGCCGAGCTGGGGA	ATCAGAGAT	300
Db	241	GGAGCTCA	CCTGCGGCCA	GAGTCCCACTA	TACTCTGCCGAGCTGGGGA	ATCAGAGAT	300
QY	301	GGAAATCA	TGTAAGCA	AGGGCCCTCGGCTA	TGAGTGTGCTGAC	TGACCCCTGTGGAC	360
Db	301	GGAAATCA	TGTAAGCA	AGGGCCCTCGGCTA	TGAGTGTGCTGAC	TGACCCCTGTGGAC	360
QY	361	GTCA	CCCCCTGTGG	CAACGGGAC	CAAGCTTGGGA	ACTTACCA	420
Db	361	GTCA	CCCCCTGTGG	CAACGGGAC	CAAGCTTGGGA	ACTTACCA	420
QY	421	CTTAA	CGGGGAC	AGCTGGAA	CCGGTGTCTCTCC	CCCAAGCCCCCGGCA	480
Db	421	CTTAA	CGGGGAC	AGCTGGAA	CCGGTGTCTCTCC	CCCAAGCCCCCGGCA	480
QY	481	GGAGCTG	GTGGCG	CAAGGCTGT	CCAAGGA	GGAAGCTGA	540
Db	481	GGAGCTG	GTGGCG	CAAGGCTGT	CCAAGGA	GGAAGCTGA	540
QY	541	CCGAGAT	CAATCA	TGTGAG	CAGACAGAT	TTCTTA	600
Db	541	CCGAGAT	CAATCA	TGTGAG	CAGACAGAT	TTCTTA	600
QY	601	GGAAGAG	AGGCTGG	CAAGCCG	CGCGAGCT	TGAAGAGCCCGGTACA	660
Db	601	GGAAGAG	AGGCTGG	CAAGCCG	CGCGAGCT	TGAAGAGCCCGGTACA	660
QY	661	GTCA	AGAC	CCGAC	CGCTGTG	TGAGATCA	720
Db	661	GTCA	AGAC	CCGAC	CGCTGTG	TGAGATCA	720

Db 661 GTGGAAGCACCGACCTGGTGTCAATCTATCGACGAGAACCGGAAAGAGCTGGAAC 720  
Qy 721 TGCACTATGSGATTTCTGAAAGGCTCTGGGGCCCGAGGTGAGTGCCTGTACAAACAGCC 780  
Db 721 TGCACTATGSGATTTCTGAAAGGCTCTGGGGCCCGAGGTGAGTGCCTGTACAAACAGCC 780  
Qy 781 CTCCGACACCGCGGAGATCATGAGAACCATCAAAATTAACAGAGCGATGCGGAAGAGCT 840  
Db 781 CTCCGACACCGCGGAGATCATGAGAACCATCAAAATTAACAGAGCGATGCGGAAGAGCT 840  
Qy 841 AATCTGTACTTCAAGAGGAGATCAAGCTGGAAGAAATGGAAGCGAATTTCTGCCA 900  
Db 841 AATCTGTACTTCAAGAGGAGATCAAGCTGGAAGAAATGGAAGCGAATTTCTGCCA 900  
Qy 901 GCGCTATGACCACTCATGAGAGGCTTGGAAAAAAAAGTGAGCGCATGAAAAACAACC 960  
Db 901 GCGCTATGACCAAGCTCATGAGAGGCTTGGAAAAAAAAGTGAGCGCATGAAAAACAACC 960  
Qy 961 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGGAGTACTACGAAAAAGCAAGTTCCTGAGAT 1020  
Db 961 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGGAGTACTACGAAAAAGCAAGTTCCTGAGAT 1020  
Qy 1021 CCGCAAGACGCGCGAGCTGCGAGAGCGCATGCAAGACAGGTGGGCGACGCGGGCAGTGG 1080  
Db 1021 CCGCAAGACGCGCGAGCTGCGAGAGCGCATGCAAGACAGGTGGGCGACGCGGGCAGTGG 1080  
Qy 1081 GCTGTCTCATGTGCGCGCGCGAGCGAGACGAGGTGTCAAGATCATGATGCGCTCTC 1140  
Db 1081 GCTGTCTCATGTGCGCGCGCGAGCGAGACGAGGTGTCAAGATCATGATGCGCTCTC 1140  
Qy 1141 AGAGCAGAGAAACTGTGAGAGACAGATGCGCAGCTGCGCTGATCCGCCATCTGTGA 1200  
Db 1141 AGAGCAGAGAAACTGTGAGAGACAGATGCGCAGCTGCGCTGATCCGCCATCTGTGA 1200  
Qy 1201 CGAAGCTGACCGACGCGCATGAAGTTCATCAACATGAAACGCGGCTTATGGCCGACCTAT 1260  
Db 1201 CGAAGCTGACCGACGCGCATGAAGTTCATCAACATGAAACGCGGCTTATGGCCGACCTAT 1260  
Qy 1261 GAAAGTGTACAAAGACCGCCAGGTATGAAACATGTGAGTGAAGAGAAAGAACTTT 1320  
Db 1261 GAAAGTGTACAAAGACCGCCAGGTATGAAACATGTGAGTGAAGAGAAAGAACTTT 1320  
Qy 1321 CCGGGAAGAAATTCAATGACAGATCCCAAGAACTTTGGCTGATGCAATCTCTGAGAG 1380  
Db 1321 CCGGGAAGAAATTCAATGACAGATCCCAAGAACTTTGGCTGATGCAATCTCTGAGAG 1380  
Qy 1381 GAAAGCAGTGGCTGAGTGGTCTCTATTACTTGACTAAGAAAGAAATGAAACTATTAA 1440  
Db 1381 GAAAGCAGTGGCTGAGTGGTCTCTATTACTTGACTAAGAAAGAAATGAAACTATTAA 1440  
Qy 1441 GAGCCTGTGAGACCGAGCTATCGGCGCGCGGCAAGAGCCAGCAGCAACACAGAGA 1500  
Db 1441 GAGCCTGTGAGACCGAGCTATCGGCGCGCGGCAAGAGCCAGCAGCAACACAGAGA 1500  
Qy 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
Db 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
Qy 1561 AGATGAGAAAGAGAAAGAAAGAGAGCGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
Db 1561 AGATGAGAAAGAGAAAGAAAGAGAGCGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
Qy 1621 CGACAAAGAAAGACTCTCTCAAGAGAAAGACAGACGACCTCAGGGGAGAGCAACAGAGA 1680  
Db 1621 CGACAAAGAAAGACTCTCTCAAGAGAAAGACAGACGACCTCAGGGGAGAGCAACAGAGA 1680  
Qy 1681 GAAAGAGGCTGTGGCTCTCAAAAGGCGCGCAAACTGCGCAACAGCGGAGAAAGCGCAAAG 1740  
Db 1681 GAAAGAGGCTGTGGCTCTCAAAAGGCGCGCAAACTGCGCAACAGCGGAGAAAGCGCAAAG 1740  
Qy 1741 CCGCATCAACCGGCTCAATGGCTTAATGAGGCCAAACAGCGAGGCGCATCAACCCCGCAGGA 1800  
Db 1741 CCGCATCAACCGGCTCAATGGCTTAATGAGGCCAAACAGCGAGGCGCATCAACCCCGCAGGA 1800

Db 1741 CCGCATCAACCGGCTCAATGGCTTAATGAGGCCAAACAGCGAGGAGGCTCATCAACCCCGCAGGA 1800  
Qy 1801 GAGCGCCGAGGTGCGCTCCATGAGACTGAAATGAGAGATTCTCGCTGAGACAGAAAGAAAT 1860  
Db 1801 GAGCGCCGAGGTGCGCTCCATGAGACTGAAATGAGAGATTCTCGCTGAGACAGAAAGAAAT 1860  
Qy 1861 GGAAGCAGCGAAAGAAAGTCTCTGGAACACGCGCGCAACTGGTGGCCATGCGCCGAT 1920  
Db 1861 GGAAGCAGCGAAAGAAAGTCTCTGGAACACGCGCGCAACTGGTGGCCATGCGCCGAT 1920  
Qy 1921 GGTGGGCTCCAAAGCTGAGTGGCAGTGTAAAGCTTTCTTCACTTCAAGAAAGAGCA 1980  
Db 1921 GGTGGGCTCCAAAGCTGAGTGGCAGTGTAAAGCTTTCTTCACTTCAAGAAAGAGCA 1980  
Qy 1981 GAACCTCGATGAGATCTTGCACAGCACAAGCTGGAATGAGAGAGAGAGAGAGAGAGAG 2040  
Db 1981 GAACCTCGATGAGATCTTGCACAGCACAAGCTGGAATGAGAGAGAGAGAGAGAGAGAG 2040  
Qy 2041 GAGAAAGAAAGAAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGT 2100  
Db 2041 GAGAAAGAAAGAAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGT 2100  
Qy 2101 GGAAGATGAGAGATGAGAGGCGTGGGGGTGAGCGGGAATGAGAGAGAGATGTTGAGAGA 2160  
Db 2101 GGAAGATGAGAGATGAGAGGCGTGGGGGTGAGCGGGAATGAGAGAGAGATGTTGAGAGA 2160  
Qy 2161 GGTCTGAACCTTTACATGCTCTGAGAAATGAGGTGCCAGAGGGGAATCAAGTGGCCAGC 2220  
Db 2161 GGTCTGAACCTTTACATGCTCTGAGAAATGAGGTGCCAGAGGGGAATCAAGTGGCCAGC 2220  
Qy 2221 CACTGTCAACACAGCTCAGACACCGAGAGCATCCCTCTCTCTCACTGAGGCGGCCAA 2280  
Db 2221 CACTGTCAACACAGCTCAGACACCGAGAGCATCCCTCTCTCTCACTGAGGCGGCCAA 2280  
Qy 2281 GGAACAAGGGGAGAAATGGGCCCCAAGCCCCAGCCACCTTGGGCGCGCGAGCGGCCACCCC 2340  
Db 2281 GGAACAAGGGGAGAAATGGGCCCCAAGCCCCAGCCACCTTGGGCGCGCGAGCGGCCACCCC 2340  
Qy 2341 AGGCCACCCACCCACCGAGAGACATCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCTC 2400  
Db 2341 AGGCCACCCACCCACCGAGAGACATCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCTC 2400  
Qy 2401 TGAAGCCACCGGAGCGCTTATGCGCCCGCACAGCACCCCATATGCGCTTGCACCTCTCC 2460  
Db 2401 TGAAGCCACCGGAGCGCTTATGCGCCCGCACAGCACCCCATATGCGCTTGCACCTCTCC 2460  
Qy 2461 TGTGGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
Db 2461 TGTGGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
Qy 2521 GAGAGAGCAGAAAGCCCCCGCGGCTGAGAGCTGAGCAGTGAACACAGGGAAGGCCGAGGA 2580  
Db 2521 GAGAGAGCAGAAAGCCCCCGCGGCTGAGAGCTGAGCAGTGAACACAGGGAAGGCCGAGGA 2580  
Qy 2581 GCCCGTCAAGAGCGAGTGCACGAGAGAAAGCGAGAGAGGCGCGGCGCAAGAGAGAGC 2640  
Db 2581 GCCCGTCAAGAGCGAGTGCACGAGAGAAAGCGAGAGAGGCGCGGCGCAAGAGAGAGC 2640  
Qy 2641 GGAAGCGCTGAGGCAAGGCGCGAGGCGCGCTCAAGGCAAGAGAGAGAGAGAGAGAGAG 2700  
Db 2641 GGAAGCGCTGAGGCAAGGCGCGAGGCGCGCTCAAGGCAAGAGAGAGAGAGAGAGAGAG 2700  
Qy 2701 CGGAGAGGCGACCACTGCGCAAGAGCTCGGGGCGCCCCCAGGACAGCGACTCCAGTGTAC 2760  
Db 2701 CGGAGAGGCGACCAAGGCAAGAGCTCGGGGCGCCCCCAGGACAGCGACTCCAGTGTAC 2760  
Qy 2761 CTGCAGTGCAAGCGAGTGAATGAGGCCGAGGCGGCGACAAAGAACTCGGCTGTCTCCC 2820  
Db 2761 CTGCAGTGCAAGCGAGTGAATGAGGCCGAGGCGGCGGCGACAAAGAACTCGGCTGTCTCCC 2820  
Qy 2821 AAGGCGAGGCTCTCTCAACCGGACTGAGAGACCCCGGGCGCAATGCTTCAACCCCAAGAGCC 2880  
Db 2821 AAGGCGAGGCTCTCTCAACCGGACTGAGAGACCCCGGGCGCAATGCTTCAACCCCAAGAGCC 2880

QY	2881	ACTGACCTTGAAAGCAGCTGAAAGACGACGCGCGTGTCCAT	CCCCCCCCCATTCAGGTCA	CCAA	2940
Dp	2881	ACTGACCTTGAAAGCAGCTGAAAGACGACGCGCGTGTCCAT	CCCCCCCCCATTCAGGTCA	CCAA	2940
QY	2941	AGTGCATGAGCCCCCGGGAGAGACGACGCTCCCA	CGAAGCAGCTCCCCAGCCCCC	3000	
Dp	2941	AGTGCATGAGCCCCCGGGAGAGACGACGCTCCCA	CGAAGCAGCTCCCCAGCCCCC	3000	
QY	3001	GCACCCGCAAAACCTGCAGCCGAGAGCGACGCTTCAG	CAGCTGGCAGACGCTCCG	3060	
Dp	3001	GCACCCGCAAAACCTGCAGCCGAGAGCGACGCTTCAG	CAGCTGGCAGACGCTCCG	3060	
QY	3061	GGGCAAGAGCAGAGAGCCCCGGACCCCCCGGCA	CAAGAGGGGCTTCGAGCGAGGGCCA	3120	
Dp	3061	GGGCAAGAGCAGAGAGCCCCGGACCCCCCGGCA	CAAGAGGGGCTTCGAGCGAGGGCCA	3120	
QY	3121	GAAGCTCGCTGGGAGACCCGCTTGCTGGA	CTTCGAGCTGCTTCGCCGCTCCG	3180	
Dp	3121	GAAGCTCGCTGGGAGACCCGCTTGCTGGA	CTTCGAGCTGCTTCGCCGCTCCG	3180	
QY	3181	TGAGGTGATCAAGGCTTCGCCGATGCGCCGAGAC	CCCTCAGCCTTCTCTACGCTCCAC	3240	
Dp	3181	TGAGGTGATCAAGGCTTCGCCGATGCGCCGAGAC	CCCTCAGCCTTCTCTACGCTCCAC	3240	
QY	3241	TGTCACCCCATCTGCCCCCTGGGCTTCAT	TGACATGCCCCGCGCTTCGCGCCGAC	3300	
Dp	3241	TGTCACCCCATCTGCCCCCTGGGCTTCAT	TGACATGCCCCGCGCTTCGCGCCGAC	3300	
QY	3301	CACCATCTCCAAACCCGCGCTCCCTCAT	CTCTCTGCGCAAGCACCAGGCTCTCGAGAG	3360	
Dp	3301	CACCATCTCCAAACCCGCGCTCCCTCAT	CTCTCTGCGCAAGCACCAGGCTCTCGAGAG	3360	
QY	3361	GCAAAATGAGTGCACATCTCCCAAGGAATGTG	GTGCACTCCACGCTCCGCTATCAGACA	3420	
Dp	3361	GCAAAATGAGTGCACATCTCCCAAGGAATGTG	GTGCACTCCACGCTCCGCTATCAGACA	3420	
QY	3421	TGCGAAGGCCCCCGGTGGGCTCTGT	CACCATGGGGCTGCCCCCTGCAATGCA	3480	
Dp	3421	TGCGAAGGCCCCCGGTGGGCTCTGT	CACCATGGGGCTGCCCCCTGCAATGCA	3480	
QY	3481	GCTGGCACCTTTCAGCGGAGTGAAGCAGAGAC	AGTGTCCCCACGGGGCCAGGCTGGGCG	3540	
Dp	3481	GCTGGCACCTTTCAGCGGAGTGAAGCAGAGAC	AGTGTCCCCACGGGGCCAGGCTGGGCG	3540	
QY	3541	ACCGGAGAGCGCTGGGGGGTGC	CCACAGGCCGAGAGGGCGTCCGTGAGAGGGACA	3600	
Dp	3541	ACCGGAGAGCGCTGGGGGGTGC	CCACAGGCCGAGAGGGCGTCCGTGAGAGGGACA	3600	
QY	3601	GGGCTCAGTTTCGGGCGGAGCAT	CACCAAGGCAATTCACGACACGCGGTGCCCTTCGGA	3660	
Dp	3601	GGGCTCAGTTTCGGGCGGAGCAT	CACCAAGGCAATTCACGACACGCGGTGCCCTTCGGA	3660	
QY	3661	CAGGCGCATCATATACGCGGGCTCAT	CAACCAAGGCAAGCAGCAGCTGATCTGTTCA	3720	
Dp	3661	CAGGCGCATCATATACGCGGGCTCAT	CAACCAAGGCAAGCAGCAGCTGATCTGTTCA	3720	
QY	3721	GGGACCATCAACAGATTCATCGGCGAGAGAC	AGCGCCGAGTGTGCTTGGACCGCGCGCGGA	3780	
Dp	3721	GGGACCATCAACAGATTCATCGGCGAGAGAC	AGCGCCGAGTGTGCTTGGACCGCGCGCGGA	3780	
QY	3781	GGACAGCTCTGCGCCAAAGGCGCACGTCAT	CTACAGAAAGGCACAGTCTTGTCTTA	3840	
Dp	3781	GGACAGCTCTGCGCCAAAGGCGCACGTCAT	CTACAGAAAGGCACAGTCTTGTCTTA	3840	
QY	3841	TGAAGGTGACATGTCTGTGACCCACAGTGCT	CCAAAGAGACGCGCAGAAAGCAGCTCAGAGC	3900	
Dp	3841	TGAAGGTGACATGTCTGTGACCCACAGTGCT	CCAAAGAGACGCGCAGAAAGCAGCTCAGAGC	3900	
QY	3901	CCCCCATGAGACGGCGCGCCGCCAACGCGAC	CTTATGACATGATGAGAGGGCGCGCTGGGGCAG	3960	
Dp	3901	CCCCCATGAGACGGCGCGCGCCGCCAACGCGAC	CTTATGACATGATGAGAGGGCGCGCTGGGGCAG	3960	

OY	3961	AGCCATCTCTCAGCCAGCATCGAAGGCTCTAATGGGCGCTGCCATTCGCCGCGAGCGACA	4020
Db	3961	AGCATCTCTCAGCCAGCATCGAAGGCTCTAATGGGCGCTGCCATTCGCCGCGAGCGACA	4020
OY	4021	CAGCCCCACACCTCAAGAGGAGCAACAATCCGGGGGCTCATCAACAAGGAGATCC	4080
Db	4021	CAGCCCCACACTTCAAGAGGAGCAACAATCCGGGGGCTCATCAACAAGGAGATCC	4080
OY	4081	TCGGTCTTACGTGAGGCAAGAAGGACTACTTGCTCGGAGGGCCAAGCTCTTAAAGC	4140
Db	4081	TCGGTCTTACGTGAGGCAAGAAGGACTACTTGCTCGGAGGGCCAAGCTCTTAAAGC	4140
OY	4141	GGAGGGACACGCTTCGCCGCCACCGCCCTCAACGGACCTTGAACGGCTTACAAGACCA	4200
Db	4141	GGAGGGACACGCTTCGCCGCCACCGCCCTCAACGGACCTTGAACGGCTTACAAGACCA	4200
OY	4201	GGCCCTGGGGCCCCCTGAAGGCTGAAGCCGGCCCAATGAAGGGCTTGATGGCCACGGTGAAGGA	4260
Db	4201	GGCCCTGGGGCCCCCTGAAGGCTGAAGCCGGCCCAATGAAGGGCTTGATGGCCACGGTGAAGGA	4260
OY	4261	GGCGGGCCGCTCATTCATGAGATCCCGCGAGAGACTCGGCAACGCGCCAGACTGCC	4320
Db	4261	GGCGGGCCGCTCATTCATGAGATCCCGCGAGAGACTCGGCAACGCGCCAGACTGCC	4320
OY	4321	CTTGGCCCGCGCGGCTCAAGAGAGGCTCATCAACGACAGGCAACCCGCTCAAGTACGA	4380
Db	4321	CTTGGCCCGCGCGGCTCAAGAGAGGCTCATCAACGACAGGCAACCCGCTCAAGTACGA	4380
OY	4381	CACCGGCGGCTTCCACCACTGGCTCMAAAGACAGATAGCTTCCTCATTCGGCAACCC	4440
Db	4381	CACCGGCGGCTTCCACCACTGGCTCMAAAGACAGATAGCTTCCTCATTCGGCAACCC	4440
OY	4441	CGGCGGAGCGTTCCTCCACCCGCTGCAACCGGCTGATGTATGAGCGCAACCGCGGAGACTGGA	4500
Db	4441	CGGCGGAGCGTTCCTCCACCCGCTGCAACCGGCTGATGTATGAGCGCAACCGCGGAGACTGGA	4500
OY	4501	ACGTGCTGTCTACGAGGAGAGCCTGAAGAGCCGGCCAGGACCGCCAGCACTTCGGGGGG	4560
Db	4501	ACGTGCTGTCTACGAGGAGAGCCTGAAGAGCCGGCCAGGACCGCCAGCACTTCGGGGGG	4560
OY	4561	CTCCATTGGCGCGCGGCCCCCGGCTATTGTGCTGACGCTGGGTAAAGCCGGGAGAAACCC	4620
Db	4561	CTCCATTGGCGCGCGGCCCCCGGCTATTGTGCTGACGCTGGGTAAAGCCGGGAGAAACCC	4620
OY	4621	CTTGAACCTTATGAGGACACGCGGGGACCTTTGCGGAGCACTCCACAGAGATTCCGCGT	4680
Db	4621	CTTGAACCTTATGAGGACACGCGGGGACCTTTGCGGAGCACTCCACAGAGATTCCGCGT	4680
OY	4681	GACCATCGGAGAGCCACGCGCGGCTTGAGAGAGGAGCACTTTGCTCAGACAAAGGATC	4740
Db	4681	GACCACCGGAGAGCCACGCGCGGCTTGAGAGAGGAGCACTTTGCTCAGACAAAGGATC	4740
OY	4741	CCAGGACCGAAAGCTGACGTGACGCTCTGTGAGATGCGCAAGTCCCAGCAAGACCGT	4800
Db	4741	CCAGGACCGAAAGCTGACGTGACGCTCTGTGAGATGCGCAAGTCCCAGCAAGACCGT	4800
OY	4801	GCCCGAGACCAACCCACACCCCATCTTGCGCTTATGAGCACTTGCTTGCGGGCGTGAATGG	4860
Db	4801	GCCCGAGACCAACCCACACCCCATCTTGCGCTTATGAGCACTTGCTTGCGGGCGTGAATGG	4860
OY	4861	CGTGAACCTGTATCGAGGACATATCCCTTGAGCTTGAACCCCACTCAATACCCCGCG	4920
Db	4861	CGTGAACCTGTATCGAGGACATATCCCTTGAGCTTGAACCCCACTCAATACCCCGCG	4920
OY	4921	CATTCCTCTGAGCGAGCGGCTGCTTACTTATCTTGCCCGACACTTGCGCCCCCAACCCAC	4980
Db	4921	CATTCCTCTGAGCGAGCGGCTGCTTACTTATCTTGCCCGACACTTGCGCCCCCAACCCAC	4980
OY	4981	CTTACCCGACCTGTATCCCAACCTTACTTACCTCGGGCTTACCCGACCGCGGGCGCTGGA	5040
Db	4981	CTTACCCGACCTGTATCCCAACCTTACTTACCTCGGGCTTACCCGACCGCGGGCGCTGGA	5040
OY	5041	GAACCGGACAGCCATCATCATGATCACTACCTCGACAGATGACACCAACGCG	5100

Db 5041 GAAACGGAGAGACATCATCAATGACTATCACTTCGACAGAGATGCAACACAAAGCCGAC 5100  
QY 5101 CACCCGATGAGCCGACGAGCTGATATGTGTGAGGGGCTCTGCCCCGAGGCTCCGCT 5160  
Db 5101 CACCCGATGAGCCGACGAGCTGATATGTGTGAGGGGCTCTGCCCCGAGGCTCCGCT 5160  
QY 5161 GGCACCTCAACTACGCTGCGGGGTCCCCGAGGACATTCGACTGTCCAAAGTCCACACT 5220  
Db 5161 GGCACCTCAACTACGCTGCGGGGTCCCCGAGGACATTCGACTGTCCAAAGTCCACACT 5220  
QY 5221 GCGTGTGCTGTGCCCCCGACACAGGACCCGACGACCCGACATGAAACCGCTTGGCTTA 5280  
Db 5221 GCGTGTGCTGTGCCCCCGACACAGGACCCGACGACCCGACATGAAACCGCTTGGCTTA 5280  
QY 5281 CTTCGCCACCGGCGCCGACGCTTTCAGACGCGGACACAGCAGCTCCCACTCTCCCGACG 5340  
Db 5281 CTTCGCCACCGGCGCCGACGCTTTCAGACGCGGACACAGCAGCTCCCACTCTCCCGACG 5340  
QY 5341 AGGTCCAAACAATTGACAAACCAACCAACCAACGTCCTGTCCGAGCGGAGCGAGACCG 5400  
Db 5341 AGGTCCAAACAATTGACAAACCAACCAACCAACGTCCTGTCCGAGCGGAGCGAGACCG 5400  
QY 5401 GGATGGAAGCGGGACCGGGATCGGGAGCGGGAAAGTCCATTCCTTCAGCTCCACAGAC 5460  
Db 5401 GGATGGAAGCGGGACCGGGATCGGGAGCGGGAAAGTCCATTCCTTCAGCTCCACAGAC 5460  
QY 5461 GGTGAGACGACGACCATCTGTGAGACCTTGATCAGAGCAGAGCAGCGGACGAGCGGACG 5520  
Db 5461 GGTGAGACGACGACCATCTGTGAGACCTTGATCAGAGCAGAGCAGCGGACGAGCGGACG 5520  
QY 5521 CAGCGCGGGGGTGGGGGACAGACGACCGCGCCGCTCCCACTCCATGCCCACAGCA 5580  
Db 5521 CAGCGCGGGGGTGGGGGACAGACGACCGCGCCGCTCCCACTCCATGCCCACAGCA 5580  
QY 5581 CTGCGCCATCTCTCCCTCGGACCCAGAGATGCCCTTCAGAGAGAACCCAGTGTCTTCAAA 5640  
Db 5581 CTGCGCCATCTCTCCCTCGGACCCAGAGATGCCCTTCAGAGAGAACCCAGTGTCTTCAAA 5640  
QY 5641 CACAGGCAATGAAGGATATCATCAACGCTGTGAGGCCAGACAGCCACGAGTCTTGAAGTC 5700  
Db 5641 CACAGGCAATGAAGGATATCATCAACGCTGTGAGGCCAGACAGCCACGAGTCTTGAAGTC 5700  
QY 5701 CACCTTCAACTCTTCAACCGGTTGCGCCAGACTGCCAATTTCCACTGCCACCCCACTGCC 5760  
Db 5701 CACCTTCAACTCTTCAACCGGTTGCGCCAGACTGCCAATTTCCACTGCCACCCCACTGCC 5760  
QY 5761 ACTGGGCGGCACTCTGATGAGGATCTTAACCTTACCTATGAGGCCGCTTGTGCTGCCAA 5820  
Db 5761 ACTGGGCGGCACTCTGATGAGGATCTTAACCTTACCTATGAGGCCGCTTGTGCTGCCAA 5820  
QY 5821 GGAAGCCCCCGGATCGCCGCGACAGAGCGGCGCCGAGAGACACCGGCAATGCTTCTCT 5880  
Db 5821 GGAAGCCCCCGGATCGCCGCGACAGAGCGGCGCCGAGAGACACCGGCAATGCTTCTCT 5880  
QY 5881 CGCCAAAGCCCCGCTCGGCTCGGAGCTGAGAGCCGCTTCTCCCGACAAAGGCTTCGGA 5940  
Db 5881 CGCCAAAGCCCCGCTCGGCTCGGAGCTGAGAGCCGCTTCTCCCGACAAAGGCTTCGGA 5940  
QY 5941 GCGCCGCGCCCTGATGCTCTGCTCTGTGACACGACCATCGCCGACCCCTGCGAA 5991  
Db 5941 GCGCCGCGCCCTGATGCTCTGCTCTGTGACACGACCATCGCCGACCCCTGCGAA 5991  
QY 5991 GCGCCGCGCCCTGATGCTCTGCTCTGTGACACGACCATCGCCGACCCCTGCGAA 5991  
Db 5991 GCGCCGCGCCCTGATGCTCTGCTCTGTGACACGACCATCGCCGACCCCTGCGAA 5991  
QY 6001 GAACTCTGACCTTCAACAAGCCGAGACCCGAGCGGCGACCACTGCTGAGGCTTCGGA 6060  
Db 6001 GAACTCTGACCTTCAACAAGCCGAGACCCGAGCGGCGACCACTGCTGAGGCTTCGGA 6060  
QY 6061 CCGCGACCGGAAAAGACTCAAGTAAACCTTTTTCATTCAGAACTGGAATCTCCGTTTC 6120  
Db 6061 CCGCGACCGGAAAAGACTCAAGTAAACCTTTTTCATTCAGAACTGGAATCTCCGTTTC 6120  
QY 6121 TCTGGTTTACCGGACGAGCTTACAGCCCGAAGGGGTGAGCCGCTGCTGTGAG 6180  
Db 6121 TCTGGTTTACCGGACGAGCTTACAGCCCGAAGGGGTGAGCCGCTGCTGTGAG 6180

Db 6112 TCTGGTTTACCGGACGAGCTTACAGCCCGAAGGGGTGAGCCGCTGCTGTGAG 6171  
QY 6181 CTACCCGAGTCTGACCCACGACAAAGGGCTCCCAAGACCTTGAAAGACTGACAAAG 6240  
Db 6181 CTACCCGAGTCTGACCCACGACAAAGGGCTCCCAAGACCTTGAAAGACTGACAAAG 6240  
QY 6241 CCACTGGAAGGGAGCTGCGGCCCAAGACGAGGCCCTGTGAACTTTGGCGGGAGGC 6300  
Db 6241 CCACTGGAAGGGAGCTGCGGCCCAAGACGAGGCCCTGTGAACTTTGGCGGGAGGC 6300  
QY 6301 GCGCCACTCCACACCTGCGCGCGCTGCGAGGAGCGCCCTGTGACAGCCGCTGCT 6360  
Db 6301 GCGCCACTCCACACCTGCGCGCGCTGCGAGGAGCGCCCTGTGACAGCCGCTGCT 6360  
QY 6361 CCAGACCGCCCGAGGGGTCAAAAGTCAACAGCGGTGTGACCTTGCGCCAGCATGAG 6420  
Db 6361 CCAGACCGCCCGAGGGGTCAAAAGTCAACAGCGGTGTGACCTTGCGCCAGCATGAG 6420  
QY 6421 TGAAGTCAATCAACAGGACTACCCGGACCACTCAACAGGACTCAGCGCACCTTGGC 6480  
Db 6421 TGAAGTCAATCAACAGGACTACCCGGACCACTCAACAGGACTCAGCGCACCTTGGC 6480  
QY 6481 GCGCCCGCTCACTCTTCCGCGGGGACAGCTGCGCCGCTGAGACCTCCGCGCCAC 6540  
Db 6481 GCGCCCGCTCACTCTTCCGCGGGGACAGCTGCGCCGCTGAGACCTCCGCGCCAC 6540  
QY 6541 CAGTACCTTCACTCCCGCCCGGACCATGTGTCGCCGCGCTGCTCCCGCCACAG 6600  
Db 6541 CAGTACCTTCACTCCCGCCCGGACCATGTGTCGCCGCGCTGCTCCCGCCACAG 6600  
QY 6601 CGAAGGGGGCAAGAGTCTTCAAGACCCAAACAAAGCTGTGTTGGGTGTGTGAGGA 6660  
Db 6601 CGAAGGGGGCAAGAGTCTTCAAGACCCAAACAAAGCTGTGTTGGGTGTGTGAGGA 6660  
QY 6661 CGGATTGGAACCTGTGTCGCCACCGAGGGGATGACGAGGACAGGGGACCTCCGAGTGC 6720  
Db 6661 CGGATTGGAACCTGTGTCGCCACCGAGGGGATGAGCGAGGACCTCCGAGTGC 6720  
QY 6721 TGTGTACCCGCTGTGTATCCGGATGAGGAAACAGACGAGCCACAGAGATGGGCTTCAA 6780  
Db 6721 TGTGTACCCGCTGTGTATCCGGATGAGGAAACAGACGAGCCACAGAGATGGGCTTCAA 6780  
QY 6781 GTCTTCAGGCAACACAGACCGCGCTTCTTCAAGAACTGACCGAGAGCACTC 6840  
Db 6781 GTCTTCAGGCAACACAGACCGCGCTTCTTCAAGAACTGACCGAGAGCACTC 6840  
QY 6841 GCGCATGTGCAAGTCCAAAGAGCAAGAGATCAACAAAGAGCTGAAACACCAACCGGAA 6900  
Db 6841 GCGCATGTGCAAGTCCAAAGAGCAAGAGATCAACAAAGAGCTGAAACACCAACCGGAA 6900  
QY 6901 TGAAGCTGAATACATATACGACGCTGGGACGAGATCTTCAATATGCCGCGCATGAC 6960  
Db 6901 TGAAGCTGAATACATATACGACGCTGGGACGAGATCTTCAATATGCCGCGCATGAC 6960  
QY 6961 CGAAACAGGCTTATGACTTATGAAAGCAGGCGGTGAGAACTGCAAGCACTCAAT 7020  
Db 6961 CGAAACAGGCTTATGACTTATGAAAGCAGGCGGTGAGAACTGCAAGCACTCAAT 7020  
QY 7021 GGGGCTGGAAGGCAATTAAGAAAGGCACTGATGGGTAAATATGACCAAGTGGGAAGATC 7080  
Db 7021 GGGGCTGGAAGGCAATTAAGAAAGGCACTGATGGGTAAATATGACCAAGTGGGAAGATC 7080  
QY 7081 CCGCGCGCTCAGCGCAATGCTTTTAAACCTCTGATGCCAGTGCAGCCTGCCGCTGC 7140  
Db 7081 CCGCGCGCTCAGCGCAATGCTTTTAAACCTCTGATGCCAGTGCAGCCTGCCGCTGC 7140  
QY 7141 TATGCCATTAACCGCTGCTGACGAGCGAGTACCACTCACCTGCGCAGGTGGCG 7200  
Db 7141 TATGCCATTAACCGCTGCTGACGAGCGAGTACCACTCACCTGCGCAGGTGGCG 7200  
QY 7201 CCGGAAGGCGCAAGGCTCTGCGACAGCCGAGAGCGGAAACCAAGTCCCGGCGCCGG 7260  
Db 7201 CCGGAAGGCGCAAGGCTCTGCGACAGCCGAGAGCGGAAACCAAGTCCCGGCGCCGG 7260  
QY 7261 CCGGAAGGCGCAAGGCTCTGCGACAGCCGAGAGCGGAAACCAAGTCCCGGCGCCGG 7261  
Db 7261 CCGGAAGGCGCAAGGCTCTGCGACAGCCGAGAGCGGAAACCAAGTCCCGGCGCCGG 7261



QY 7261 CTTGGATCTGGGAGCCGGCCACTCTGTCTCTGAGTCACTCGAGGAGAGACTGCAA 7320  
 DB CTTGGATCTGGGAGCCGGCCACTCTGTCTCTGAGTCACTCGAGGAGAGACTGCAA 7311  
 QY 7321 CGGCGGAGCGCGCTCAACCAACCGCTGTGGAGAGACAGGCGCTGTCTCGAGTTCCAC 7380  
 DB 7312 CGGCGGAGCGCGCTCAACCAACCGCTGTGGAGAGACAGGCGCTGTCTCGAGTTCCAC 7371  
 QY 7381 GGCATTCCTTCAACCCCTGATCAATGGGCTGAGGGGGGTGATAGGCTTCCACAC 7440  
 DB 7372 GGCATTCCTTCAACCCCTGATCAATGGGCTGAGGGGGGTGATAGGCTTCCACAC 7431  
 QY 7441 CCCACCGGCGCTCCCGCGGAGCGGCGCTCTGCTGGCGCCCAACAGCTGGAGACA 7500  
 DB 7432 CCCACCGGCGCTCCCGCGGAGCGGCGCTCTGCTGGCGCCCAACAGCTGGAGACA 7491  
 QY 7501 GGAAGCCCAACCACTCTCTCTGCAAGTACAGACACTCTCGACAGCGAGTACTCAG 7560  
 DB 7492 GGAAGCCCAACCACTCTCTCTGCAAGTACAGACACTCTCGACAGCGAGTACTCAG 7551  
 QY 7561 AACAGGGCGGGGGGGGGCGGGCGGTGUBSPERLUCTCAAGTCCAGCGACCAAGGAA 7620  
 DB 7552 AACAGGGCGGGGGGGGGCGGGCGGTGUBSPERLUCTCAAGTCCAGCGACCAAGGAA 7599  
 QY 7621 CGGCGCTGAGAGCGGGCGGCGCTGCGCACTCCCAACCAAGAGAGAGCCCTGAGTC 7680  
 DB 7600 CGGCGCTGAGAGCGGGCGGCGCTGCGCACTCCCAACCAAGAGAGAGCCCTGAGTC 7659  
 QY 7681 CGCCTGCGCTCTCATCATCTGTCTGCTCAAGAGCGGCACTCTGCTGTAAAGCTT 7740  
 DB 7660 CGCCTGCGCTCTCATCATCTGTCTGCTCAAGAGCGGCACTCTGCTGTAAAGCTT 7719  
 QY 7741 AACTAAGACTCCCGCGGGCGGCTGTGAGACCTTACTCAGGGGATGTTAACCTG 7800  
 DB 7720 AACTAAGACTCCCGCGGGCGGCTGTGAGACCTTACTCAGGGGATGTTAACCTG 7779  
 QY 7801 GTGCTGAGAGAGGAGAGAGGAGCGGAGAGGAGAGCGGAGCGGCTGTGAGAGCCAC 7860  
 DB 7780 GTGCTGAGAGAGGAGAGAGGAGCGGAGAGGAGAGCGGAGCGGCTGTGAGAGCCAC 7839  
 QY 7861 ACACAGGCGGCGGAGAGGAGCGGAGAGGAGAGCGGAGCGGAGCGGAGCGGAGCGGAG 7920  
 DB 7840 ACACAGGCGGCGGAGAGGAGCGGAGAGGAGAGCGGAGCGGAGCGGAGCGGAGCGGAG 7899  
 QY 7921 TGGCTTCCCGGAGATGATTTGGAACCAAGTCTTAACTGAGCTCGAGCGCCCGGCTT 7980  
 DB 7900 TGGCTTCCCGGAGATGATTTGGAACCAAGTCTTAACTGAGCTCGAGCGCCCGGCTT 7959  
 QY 7981 CCTCGGCTCCCATCTCGGCTTACGAGCTTGAAGAGAGAGCGGAGCGGCTTGAAGAG 8040  
 DB 7960 CCTCGGCTCCCATCTCGGCTTACGAGCTTGAAGAGAGAGCGGAGCGGCTTGAAGAG 8019  
 QY 8041 CCAAGTGTGCTGTCTGCTGCTCCCAAGAGTGTCCCAAGAGATTTCTGGAAACCA 8100  
 DB 8020 CCAAGTGTGCTGTCTGCTGCTCCCAAGAGTGTCCCAAGAGATTTCTGGAAACCA 8079  
 QY 8101 GTACAGGCGGAGTGTGAGCAAAAGGCGAGGCTGTGGGGGAGAGAGATGCTCCGA 8160  
 DB 8080 GTACAGGCGGAGTGTGAGCAAAAGGCGAGGCTGTGGGGGAGAGAGATGCTCCGA 8139  
 QY 8161 GGAAGTGAATGTTTTTTTCAACATGTTGCGGAGCGGAGGAGAGAGAGAGATGTA 8220  
 DB 8140 GGAAGTGAATGTTTTTTTCAACATGTTGCGGAGCGGAGGAGAGAGAGATGTA 8199  
 QY 8221 AATGATGTGTGTTTACAGGATATATTTTGAATACCTTCAATGAATTAATTCAGATGT 8280  
 DB 8200 AATGATGTGTGTTTACAGGATATATTTTGAATACCTTCAATGAATTAATTCAGATGT 8259  
 QY 8281 TTAAGCAAGAGAGATTAACCAAGTATTAAGTGTGCTGTGCTTTTATCTGTGCTTACCG 8340  
 DB 8260 TTAAGCAAGAGAGATTAACCAAGTATTAAGTGTGCTGTGCTTTTATCTGTGCTTACCG 8319

QY 8341 TTCAGAGGCGGTGTGAGAGGCGGACAGTGTGTAGCCCATCACTCGAGAGACCAAGGGGCG 8400  
 DB 8320 TTCAGAGGCGGTGTGAGAGGCGGACAGTGTGTAGCCCATCACTCGAGAGACCAAGGGGCG 8379  
 QY 8401 GGGGACTGTCTCG-TACGCGCCGCTGTGTCTTCCCTTCTTCTTGGGAGATGA 8459  
 DB 8380 GGGGACTGTCTCG-TACGCGCCGCTGTGTCTTCCCTTCTTCTTGGGAGATGA 8439  
 QY 8460 ATTGATGCGTATTTCTGTGAGCGGCAATTGCGAGGAGGTGTATTTCTGATTACAG 8519  
 DB 8440 ATTGATGCGTATTTCTGTGAGCGGCAATTGCGAGGAGGTGTATTTCTGATTACAG 8499  
 QY 8520 ACGTGTCTTAATTAATAAAGCGAATTATTAATCTCA 8553  
 DB 8500 ACGTGTCTTAATTAATAAAGCGAATTATTAATCTCA 8533  
 RESULT 6  
 ACN39603  
 ID ACN39603 standard; cDNA; 8533 BP.  
 AC  
 XX ACN39603;  
 DT  
 XX 18-NOV-2004 (first entry)  
 DE  
 XX Tumour-associated antigenic target (TAT) cDNA DNA325805, SEQ ID NO:3856.  
 KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;  
 KW tumour; diagnosis; cell proliferative disorder; breast cancer;  
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;  
 KW central nervous system cancer; bladder cancer; pancreatic cancer;  
 KW cervical cancer; melanoma; leukaemia; hybridisation probe;  
 KW chromosome identification; chromosome mapping; gene mapping;  
 KW gene therapy; cytostatic; gene; ss.  
 OS  
 XX Homo sapiens.  
 XX  
 XX MO2004030615-A2.  
 XX  
 PD 15-APR-2004.  
 PF 29-SEP-2003; 2003WO-US028547.  
 XX  
 PR 02-OCT-2002; 2002US-0414971P.  
 XX  
 PA (GENTH ) GENENTECH INC.  
 XX  
 XX Wu TD, Zhang Z, Zhou Y;  
 PI  
 XX  
 DR WPI; 2004-347921/32.  
 XX  
 PT New tumor-associated antigenic target polypeptides and nucleic acids,  
 PT useful in preparing a medicament for treating or detecting a  
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
 PT prostate cancer or tumor.  
 XX  
 XX  
 PS Claim 1: SEQ ID NO 3856; 7273bp; English.  
 XX  
 CC The invention relates to human tumour-associated antigenic target (TAT)  
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are  
 CC overexpressed in cancer tissues compared to normal tissues, and may thus  
 CC serve as effective targets for the diagnosis and treatment of cancer in  
 CC mammals. The invention also relates to nucleic acid and polypeptide  
 CC sequences at least 80% identical to the TAT nucleic acids and  
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic  
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic  
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a  
 CC TAT polypeptide; and methods and compositions for the treatment or  
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,  
 CC antibodies, antagonists, binding molecules and compositions are useful  
 CC for diagnosing or treating a cell proliferative disorder associated with  
 CC increased TAT expression, particularly cancers such as breast cancer,  
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder



CC cancer, pancreatic cancer, cervical cancer, cancers of the central  
CC nervous system, melanoma and leukemia. TAT nucleic acids may further be  
CC used as hybridisation probes, in chromosome and gene mapping, in  
CC chromosome identification and in gene therapy. The present sequence  
CC represents a TAT nucleic acid of the invention  
XX

Sequence 8533 BP; 1836 A; 3034 C; 2528 G; 1135 T; 0 U; 0 Other;

Query Match 98.3%; Score 8423.4; DB 13; Length 8533;  
Best Local Similarity 99.3%; Pred. No. 0;

Matches 8493; Conservative 3; Mismatches 36; Indels 22; Gaps 3;

QY 1 CATGTCGGGCTTCAACAAGCTTGTGGCACAGAGTGGAGGGCCACTAGAGCCCGCTACCC 60  
DB 1 CATGTCGGGATCCACAAGCTGTGGCACAGAGTGGAGGGCCACTAGAGCCCGCTACCC 60  
QY 61 GCCCAGAGCCCTTCTTACCGATGAGATGCCCGGAGAGCACAGAGATGGTGGGCTCCT 120  
DB 61 GCCCAGAGCCCTTCTTACCGATGAGATGCCCGGAGAGCACAGAGATGGTGGGCTCCT 120  
QY 121 GGAGTACAGACAGACACTCCCGGACTATGCTCCCACTGTGCGGGGCTCCATATCCA 180  
DB 121 GGAGTACAGACAGACACTCCCGGACTATGCTCCCACTGTGCGGGGCTCCATATCCA 180  
QY 181 GCCCAGAGGGCGGAGGCTCCTGCTGTGTAGATTCCAGCCCGGAGATGAACGGTCCCA 240  
DB 181 GCCCAGAGGGCGGAGGCTCCTGCTGTGTAGATTCCAGCCCGGAGATGAACGGTCCCA 240  
QY 241 GGAGCTCCACCTGGGGCGAGATGCCACATACCTGCGCCGAGCTGGGGAACTCCAGAGT 300  
DB 241 GGAGCTCCACCTGGGGCGAGATGCCACATACCTGCGCCGAGCTGGGGAACTCCAGAGT 300  
QY 301 GGAGTTTCATTGAAACCAAGCGCTCTCGGCTAGAGCTGTGCTGACCCCTGCTCGAC 360  
DB 301 GGAGTTTCATTGAAACCAAGCGCTCTCGGCTAGAGCTGTGCTGACCCCTGCTCGAC 360  
QY 361 GTACACCCCTGCTGGCCAGCGGCTGCGGGATCTGAAGCTTCACCAAGAACCGTAG 420  
DB 361 GTACACCCCTGCTGGCCAGCGGCTGCGGGATCTGAAGCTTCACCAAGAACCGTAG 420  
QY 421 CCTGACGGGCAAGCTGGAACCGGTGCTCTCCCGGAGCCCGGCACTGAGCCCTGAGCT 480  
DB 421 CCTGACGGGCAAGCTGGAACCGGTGCTCTCCCGGAGCCCGGCACTGAGCCCTGAGCT 480  
QY 481 GGAGCTGTGTCCCGCAGCGCTGTCCAAAGGAGGCTGATCCAGAACATGGAACGGTGA 540  
DB 481 GGAGCTGTGTCCCGCAGCGCTGTCCAAAGGAGGCTGATCCAGAACATGGAACGGTGA 540  
QY 541 CCGAGAGATCACCATGTGAGACAGACAGATCTCTAAGCTGAAGAAAGAACAGCAAGCT 600  
DB 541 CCGAGAGATCACCATGTGAGACAGACAGATCTCTAAGCTGAAGAAAGAACAGCAAGCT 600  
QY 601 GGAGGAGAGAGGTGSCAAGCGGCGCGAGCTGAGAAAGCCCGTGTACCGCGCCCATGA 660  
DB 601 GGAGGAGAGAGGTGSCAAGCGGCGCGAGCTGAGAAAGCCCGTGTACCGCGCCCATGA 660  
QY 661 GTGCAAGACCGGCAAGCTGTGTGCAAGATCATCAACAGAGAACCGGAAAGAGCTGAAG 720  
DB 661 GTGCAAGACCGGCAAGCTGTGTGCAAGATCATCAACAGAGAACCGGAAAGAGCTGAAG 720  
QY 721 TGACATCATGTGATTTGGAAGGCTTGAGGGCCCAAGGTGAGCTGCGCTGTACCAAGCC 780  
DB 721 TGACATCATGTGATTTGGAAGGCTTGAGGGCCCAAGGTGAGCTGCGCTGTACCAAGCC 780  
QY 781 CTCGCAACCCGGCGATCATGAGAACATCAAAATTAACCAAGGAGTGGGAAAGAACT 840  
DB 781 CTCGCAACCCGGCGATCATGAGAACATCAAAATTAACCAAGGAGTGGGAAAGAACT 840  
QY 841 AATCTTGATCTTCAAGAGAGGAATCAGCTCGAGAAACAATGAGAGAGAACTTTCGCA 900  
DB 841 AATCTTGATCTTCAAGAGAGGAATCAGCTCGAGAAACAATGAGAGAGAACTTTCGCA 900  
QY 901 GCGCTATGACCAAGCTCATGAGGCTTGAGAAAAAAGGTGAGCGCATGAAAAACAACC 960  
DB 901 GCGCTATGACCAAGCTCATGAGGCTTGAGAAAAAAGGTGAGCGCATGAAAAACAACC 960

DB 901 GCGCTATGACCAAGCTCATGAGGCTTGAGAGAAAGGTGAGCGCATGAAAAACAACC 960  
QY 961 GCGCGCGCGGCGCAAGAGAGCAAGTGCAGAGTACTACGAAGAAAGCAATTCCTGAGAT 1020  
DB 961 GCGCGCGCGGCGCAAGAGAGCAAGTGCAGAGTACTACGAAGAAAGCAATTCCTGAGAT 1020  
QY 1021 CCGCAAGACGCGAGCTGCAGAGACGATGACAGAGAGTGGGCGACGCGGAGAGTGG 1080  
DB 1021 CCGCAAGACGCGAGCTGCAGAGACGATGACAGAGAGTGGGCGACGCGGAGAGTGG 1080  
QY 1081 GCTGTCCATGTGTGCGCGCGCGAGAGACAGAGGTGTCAAGATCATGATGAGCTCTC 1140  
DB 1081 GCTGTCCATGTGTGCGCGCGCGAGAGACAGAGGTGTCAAGATCATGATGAGCTCTC 1140  
QY 1141 AGAGCAGAGAAACCTGGAAGACAGATGCGCAGCTGCGCGATCCGCGCCATGCTGTGA 1200  
DB 1141 AGAGCAGAGAAACCTGGAAGACAGATGCGCAGCTGCGCGATGCGCGCCATGCTGTGA 1200  
QY 1201 CGACGCTGACAGACGCGCATCAAGTTTCATCAACATGAACGGGCTTATGGCGACCCGAT 1260  
DB 1201 CGACGCTGACAGACGCGCATCAAGTTTCATCAACATGAACGGGCTTATGGCGACCCGAT 1260  
QY 1261 GAAGGTGTACAAAGACCCCGAGTTCATGAACATGTGAGTGAAGAGAGAAAGCACTT 1320  
DB 1261 GAAGGTGTACAAAGACCCCGAGTTCATGAACATGTGAGTGAAGAGAGAAAGCACTT 1320  
QY 1321 CCGGAGAAAGTTTCATGACAGATCCCAAGAACTTTGGGCTGATCCGATCTTGGAGAG 1380  
DB 1321 CCGGAGAAAGTTTCATGACAGATCCCAAGAACTTTGGGCTGATCCGATCTTGGAGAG 1380  
QY 1381 GAAGACAGTGTGATGATGTCCTCTTACTTACTGCTGACTGAAGAAATGAACTATTA 1440  
DB 1381 GAAGACAGTGTGATGATGTCCTCTTACTTACTGCTGACTGAAGAAATGAACTATTA 1440  
QY 1441 GAGCTGTGTGAGACGAGACTTTCGCGCGCGGCAAGAGCCAGACAGCAACAGACGA 1500  
DB 1441 GAGCTGTGTGAGACGAGACTTTCGCGCGCGGCAAGAGCCAGACAGCAACAGACGA 1500  
QY 1501 GCAGCAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAA 1560  
DB 1501 GCAGCAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAA 1560  
QY 1561 AGATGAGAAAGAGAAAGAAAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
DB 1561 AGATGAGAAAGAGAAAGAAAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
QY 1621 CGACAAAGAAAGACTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
DB 1621 CGACAAAGAAAGACTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
QY 1681 GAAAGAGGCTGTGGCTTCAAGAGCCGCAAAACTGCCCAAGACGAGAGAGAGAGAG 1740  
DB 1681 GAAAGAGGCTGTGGCTTCAAGAGCCGCAAAACTGCCCAAGACGAGAGAGAGAGAG 1740  
QY 1741 CCGCATCAACCCGCTCAATGAGTAAATGAGGCCAAGACGAGAGAGGCAATCACCCCA 1800  
DB 1741 CCGCATCAACCCGCTCAATGAGTAAATGAGGCCAAGACGAGAGAGGCAATCACCCCA 1800  
QY 1801 GAGCGCCGAGCTGTGCTCATGAGCTGATGAGATTTCTGCTGTGACAGAAAGAAAT 1860  
DB 1801 GAGCGCCGAGCTGTGCTCATGAGCTGATGAGATTTCTGCTGTGACAGAAAGAAAT 1860  
QY 1861 GGAAGACGCGCAAGAAAGGTCTCTGGAACAGAGCGGCAAGTGTGCGGCAATCGCCGAT 1920  
DB 1861 GGAAGACGCGCAAGAAAGGTCTCTGGAACAGAGCGGCAAGTGTGCGGCAATCGCCGAT 1920  
QY 1921 GGTGGCTCTCAAGACTGTGTGCAAGTGAAGACTTCTCACTTCAACTCAAGAAAGAGCA 1980  
DB 1921 GGTGGCTCTCAAGACTGTGTGCAAGTGAAGACTTCTCACTTCAACTCAAGAAAGAGCA 1980  
QY 1981 GAACTCGATGAGATCTTTCAGACAGACAGCTGAAGATGAGAGAGAGAGAGAGCGCG 2040  
DB 1981 GAACTCGATGAGATCTTTCAGACAGACAGCTGAAGATGAGAGAGAGAGAGAGAGCGCG 2040

[illegible]

QY	3121	GAAGCTGCTGGGAGACCCCTTCTGCTGAGACTTCGGGCTGCTTCCTCCCGTGGCTCCCGC	3180
Db	3121	GAAGCTGCTGGGAGACCCCTTCTGAGACTTCGGGCTGCTTCCTCCCGTGGCTCCCGC	3180
QY	3181	TGAGGTGATCAAGGCTCCCGGAGATGGCCGGGAGCCGTCAGGCTTCCTAGCTCCAGC	3240
Db	3181	TGAGGTGATCAAGGCTCCCGGAGATGGCCGGGAGCCGTCAGGCTTCCTAGCTCCAGC	3240
QY	3241	TGTCACCACTGCGGCTGGGCTCTCCATGACATCTCCGCGGCTGCTCTGCGCGCCAGC	3300
Db	3241	TGTCACCACTGCGGCTGGGCTCTCCATGACATCTCCGCGGCTGCTCTGCGCGCCAGC	3300
QY	3301	CACCATCTCCAAACCGGCTCTCCCTCATCTCTCTGCGAAGCACCCAGGCTCTCGAGAG	3360
Db	3301	CACCATCTCCAAACCGGCTCTCCCTCATCTCTCTGCGAAGCACCCAGGCTCTCGAGAG	3360
QY	3361	GCAATATAGTGCATCTCCCAAGGAGATGTGGTCCACTCCACTCCCGTATCTCAGAGCA	3420
Db	3361	GCAATATAGTGCATCTCCCAAGGAGATGTGGTCCACTCCACTCCCGTATCTCAGAGCA	3420
QY	3421	TGCCAAGGCCCCGGTGGGCCCCCTGCACCATGAGGGCTGCCCTGGCCATGAGACCCCAAAA	3480
Db	3421	TGCCAAGGCCCCGGTGGGCCCCCTGCACCATGAGGGCTGCCCTGGCCATGAGACCCCAAAA	3480
QY	3481	GCTGCGACCTCTTACGCGGAGTGAAGCAGAGCAGCTGTCCCAAGGGGCGCAGGCTGGGCT	3540
Db	3481	GCTGCGACCTCTTACGCGGAGTGAAGCAGAGCAGCTGTCCCAAGGGGCGCAGGCTGGGCT	3540
QY	3541	ACCAGAGAGCTTGGGGGGTGCCTCACAAGCCAGAGAGGCTCCGTGCTGAGAGGAGACGCTT	3600
Db	3541	ACCAGAGAGCTTGGGGGGTGCCTCACAAGCCAGAGAGGCTCCGTGCTGAGAGGAGACGCTT	3600
QY	3601	GGGCTCAGTTCCGGGCGGAGAGCATCCAAAGGATTCAGACACAGGGTGCCCTGGGA	3660
Db	3601	GGGCTCAGTTCCGGGCGGAGAGCATCCAAAGGATTCAGACACAGGGTGCCCTGGGA	3660
QY	3661	CAGGCGCATCATATACGCGGCTCATCAACCACGAGCAGCAGCTGACTCTGTCAA	3720
Db	3661	CAGGCGCATCATATACGCGGCTCATCAACCACGAGCAGCAGCTGACTCTGTCAA	3720
QY	3721	GGGCAACATCACAGATCATCGGCGAGAGCAGCCGAGTGGCTTGGACCGCGGCGGGGA	3780
Db	3721	GGGCAACATCACAGATCATCGGCGAGAGCAGCCGAGTGGCTTGGACCGCGGCGGGGA	3780
QY	3781	GGAACAGCTGCCCCAAGGCGCACGTCATCTAGAGGAGAAAGGCGCAAGCTTGTCTTA	3840
Db	3781	GGAACAGCTGCCCCAAGGCGCACGTCATCTAGAGGAGAAAGGCGCAAGCTTGTCTTA	3840
QY	3841	TGAGGGTGGACATGTCGTGAGCCACAGTCTCCAAAGAGAGACGCGAGAGAGCTCAGAGCC	3900
Db	3841	TGAGGGTGGACATGTCGTGAGCCACAGTCTCCAAAGAGAGACGCGAGAGAGCTCAGAGCC	3900
QY	3901	CCCCCATGAGACGGCGCGCCCCAAGCGCACCTATGACATGATGAGAGGCGCGGTGGGAG	3960
Db	3901	CCCCCATGAGACGGCGCGCCCCAAGCGCACCTATGACATGATGAGAGGCGCGGTGGGAG	3960
QY	3961	AGCCATCTCTCAGCGCAGATCGAAGGTCTCATGAGGCGGTGCATCTCCCGCGAGGCACA	4020
Db	3961	AGCCATCTCTCAGCGCAGATCGAAGGTCTCATGAGGCGGTGCATCTCCCGCGAGGCACA	4020
QY	4021	CAGCCCCACCACTCAAGAGAGCAGACCATCGGCGGGTGCATCAACAAGGGATCCC	4080
Db	4021	CAGCCCCACCACTCAAGAGAGCAGACCATCGGCGGGTGCATCAACAAGGGATCCC	4080
QY	4081	TGCGTCTTACGTGAGGACAGAGAGCTACTTGCGTGGGAGGCGCAAGTCTTAAAGCG	4140
Db	4081	TGCGTCTTACGTGAGGACAGAGAGCTACTTGCGTGGGAGGCGCAAGTCTTAAAGCG	4140
QY	4141	GGAAGGCAAGCTTCCGCCCCCAACGCGCTCACAAGGAGCTTGAACGAGGCTTACAAAGAGCA	4200
Db	4141	GGAAGGCAAGCTTCCGCCCCCAACGCGCTCACAAGGAGCTTGAACGAGGCTTACAAAGAGCA	4200

Qy 4201 GGCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCTGTGGCCACGGTGAAGA 4260  
| | | | |  
Db 4201 GGCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCTGTGGCCACGGTGAAGA 4260  
| | | | |  
Qy 4281 GGGGGGGGGCTTCATTCATGAGATCCCGGCGAGAGGCTGCCGACACGCCGAGCTGCC 4320  
| | | | |  
Db 4281 GGGGGGGGGCTTCATTCATGAGATCCCGGCGAGAGGCTGCCGACACGCCGAGCTGCC 4320  
| | | | |  
Qy 4321 CTGGCCCCCGGGCGCTCAAGAGGGCTCATCAAGCAAGGACCCCGCTCAAGTACGA 4380  
| | | | |  
Db 4321 CTGGCCCCCGGGCGCTCAAGAGGGCTCATCAAGCAAGGACCCCGCTCAAGTACGA 4380  
| | | | |  
Qy 4381 CACCGGCGGCTCAACCATCTGGCTCCAAAAAGCAGACGTACGCTCTCATCGGACGCC 4440  
| | | | |  
Db 4381 CACCGGCGGCTCAACCATCTGGCTCCAAAAAGCAGACGTACGCTCTCATCGGACGCC 4440  
| | | | |  
Qy 4441 CGGCGGAGCGTTTCCACCCCTGTCACCCGCTGGATGTGATGGCCGACGCCGGGCACTGGA 4500  
| | | | |  
Db 4441 CGGCGGAGCGTTTCCACCCGCTGGATGTGATGGCCGACGCCGGGCACTGGA 4500  
| | | | |  
Qy 4501 ACGTCCCTGCTACAGAGAGGCTGAAAGCCGGCCAGGAGCCGCGACAGAGCTCGGGGG 4560  
| | | | |  
Db 4501 ACGTCCCTGCTACAGAGAGGCTGAAAGCCGGCCAGGAGCCGCGACAGAGCTCGGGGG 4560  
| | | | |  
Qy 4561 CTCCATTGCGCGGGCGCCCGGTCATTGTGCTGAGCTGGGTAAAGCCGGCAGAGCCC 4620  
| | | | |  
Db 4561 CTCCATTGCGCGGGCGCCCGGTCATTGTGCTGAGCTGGGTAAAGCCGGCAGAGCCC 4620  
| | | | |  
Qy 4621 CCTGACCTATGAGGACACAGGGGCAACCTTTGCCGGCCACTTCCACAGAGTTCCCGCT 4680  
| | | | |  
Db 4621 CCTGACCTATGAGGACACAGGGGCAACCTTTGCCGGCCACTTCCACAGAGTTCCCGCT 4680  
| | | | |  
Qy 4681 GACCATGCGGGAGCCCAAGCCGCGCTGCAAGAGGGGCAAGCTTTGTCAGCAAGCATC 4740  
| | | | |  
Db 4681 GACCATGCGGGAGCCCAAGCCGCGCTGCAAGAGGGGCAAGCTTTGTCAGCAAGCATC 4740  
| | | | |  
Qy 4741 CCAAGACCGAAAGCTGACGCTGACGCTCTGTGAGATGCGCAAGTCCCGCAGACCGT 4800  
| | | | |  
Db 4741 CCAAGACCGAAAGCTGACGCTGACGCTCTGTGAGATGCGCAAGTCCCGCAGACCGT 4800  
| | | | |  
Qy 4801 GCCCGAAGACCCACACACCCCATCTCGCCCTATGAGCACTGTCCGGGGGTGAGTGG 4860  
| | | | |  
Db 4801 GCCCGAAGACCCACACACCCCATCTCGCCCTATGAGCACTGTCCGGGGGTGAGTGG 4860  
| | | | |  
Qy 4861 CGTGAACCTGTATCGCAGGACATCCCCCTGSCCTTTCGACCCCACTCCATACCCCGGG 4920  
| | | | |  
Db 4861 CGTGAACCTGTATCGCAGGACATCCCCCTGSCCTTTCGACCCCACTCCATACCCCGGG 4920  
| | | | |  
Qy 4921 CATCCCTGTGACGAGCGCGCTGCTACTACTGTGCCCGGACACTGTGCCCAACCCAC 4980  
| | | | |  
Db 4921 CATCCCTGTGACGAGCGCGCTGCTACTACTGTGCCCGGACACTGTGCCCAACCCAC 4980  
| | | | |  
Qy 4981 CTACCCCGAACCTGTATCCCACTTACTCTATCCCGGGCTTACCCCGACGCGGGCGCTGGA 5040  
| | | | |  
Db 4981 CTACCCCGAACCTGTATCCCACTTACTCTATCCCGGGCTTACCCCGACGCGGGCGCTGGA 5040  
| | | | |  
Qy 5041 GAAACCGGAGACCATCATATGACTATCACTCACTCGCAGAGATGACACAAACAGCG 5100  
| | | | |  
Db 5041 GAAACCGGAGACCATCATATGACTATCACTCACTCGCAGAGATGACACAAAGCGCG 5100  
| | | | |  
Qy 5101 CACCGCATGTGCCAGGAGCTGATATGTGAGGGGCTTGTGCCCGGAGTCTCGGT 5160  
| | | | |  
Db 5101 CACCGCATGTGCCAGGAGCTGATATGTGAGGGGCTTGTGCCCGGAGTCTCGGT 5160  
| | | | |  
Qy 5161 GGCACTCACTATCGCTGGGGGTCCCGAGGATTCATCGACTGTTCGCAAGTGCACACT 5220  
| | | | |  
Db 5161 GGCACTCACTATCGCTGGGGGTCCCGAGGATTCATCGACTGTTCGCAAGTGCACACT 5220  
| | | | |  
Qy 5221 GCGTGTGCTGTCGTCGCGAGACAGGAGCCCGCAGCACCCTGATGGAACCGGCTTGGCTTA 5280  
| | | | |  
Db 5221 GCGTGTGCTGTCGTCGCGAGACAGGAGCCCGCAGCACCCTGATGGAACCGGCTTGGCTTA 5280  
| | | | |  
Qy 5281 CCTCCCAACCGGCGCCCAAGCCTTTCAGAGCGGCAAGCAAGCTTCCCACTTCCCAAG 5340  
| | | | |

Db 5281 | | | | | 5340  
| | | | |  
Qy 5341 AGTCCCAACACACTTGAACAAAACCAACACAGCTCTCTGTCCGAGCGGGAGCGAGACG 5400  
| | | | |  
Db 5341 AGTCCCAACACACTTGAACAAAACCAACACAGCTCTCTGTCCGAGCGGGAGCGAGACG 5400  
| | | | |  
Qy 5401 GGATCGAGAGCGGAGCCGGAGTCGGGAGCGGGAAAAGTCCATCTCAGCTCCACCAAGAC 5460  
| | | | |  
Db 5401 GGATCGAGAGCGGAGCCGGAGTCGGGAGCGGGAAAAGTCCATCTCAGCTCCACCAAGAC 5460  
| | | | |  
Qy 5461 GGTGAGGACCGACCCCATCTGTGAGACTTGTGTACAGAGCAGACAGCGGAGCGGAG 5520  
| | | | |  
Db 5461 GGTGAGGACCGACCCCATCTGTGAGACTTGTGTACAGAGCAGACAGCGGAGCGGAG 5511  
| | | | |  
Qy 5521 CAGCGGGGGGGTGGGGGAGAGAGCGCGCGCGCTCCCACTCCCATGCCCCACAGCA 5580  
| | | | |  
Db 5521 CAGCGGGGGGGTGGGGGAGAGAGCGCGCGCGCTCCCACTCCCATGCCCCACAGCA 5571  
| | | | |  
Qy 5581 CTGCGCCATCTTCCCTCGGACCCAGAGATGCCCTCCAGACAGACCCAGTGTCTTACAA 5640  
| | | | |  
Db 5572 CTGCGCCATCTTCCCTCGGACCCAGAGATGCCCTCCAGACAGAGCCCAATGTGCTTACAA 5631  
| | | | |  
Qy 5641 CACAGGCAATGAAGGTATCATCAACGCTGTGAGCCCAAGCAAGCCCAAGTCTGAGGTC 5700  
| | | | |  
Db 5632 CACAGGCAATGAAGGTATCATCAACGCTGTGAGCCCAAGCAAGCCCAAGTCTGAGGTC 5691  
| | | | |  
Qy 5701 CACCTCCACTCTCTACCCGTTGCGCCAGCTGCCACATTTCCACCTGACCCCACTGCC 5760  
| | | | |  
Db 5692 CACCTCCACTCTCTACCCGTTGCGCCAGCTGCCACATTTCCACCTGACCCCACTGCC 5751  
| | | | |  
Qy 5761 ACTGGGCGGCAACCTCGATGGGGGTCTACCTTACCTCATGAGAGCCGCTTGTGCTCCAA 5820  
| | | | |  
Db 5752 ACTGGGCGGCAACCTCGATGGGGGTCTACCTTACCTCATGAGAGCCGCTTGTGCTCCAA 5811  
| | | | |  
Qy 5821 GAGAGCCCCCGGGTGGCCCCGAGAGCGGCCGAGACACACCGGCTATGCTTCTT 5880  
| | | | |  
Db 5812 GAGAGCCCCCGGGTGGCCCCGAGAGCGGCCGAGACACACCGGCTATGCTTCTT 5871  
| | | | |  
Qy 5881 CGGCAAGCCCCGAGCGGCTCGGGCTGGAGCCGCGCTCTCCGCCCAAGAGGCTCGGA 5940  
| | | | |  
Db 5872 CGGCAAGCCCCGAGCGGCTCGGGCTGGAGCCGCGCTCTCCGCCCAAGAGGCTCGGA 5931  
| | | | |  
Qy 5941 GCGCGGCGCCCTAGTGTCTCTGTGCTGTGAGCAGCCACATGCGCCGACCCCTCGCA 6000  
| | | | |  
Db 5932 GCGCGGCGCCCTAGTGTCTCTGTGCTGTGAGCAGCCAGCCCGCGGCACTGTGCTCGGA 5991  
| | | | |  
Qy 6001 GAACTCGACCTTCAACAGCCAGCCGAGCCCGCGGCGCACTGTGCTGAGCTCGGA 6060  
| | | | |  
Db 5992 GAACTCGACCTTCAACAGCCAGCCGAGCCCGCGGCGCACTGTGCTGAGCTCGGA 6051  
| | | | |  
Qy 6061 CCGGACCGGGAAAAAGTCAAAAGTAAACCTTTTCCATCCAGAGAACTGGAATCCGTTTC 6120  
| | | | |  
Db 6052 CCGGACCGGGAAAAAGTCAAAAGTAAACCTTTTCCATCCAGAGAACTGGAATCCGTTTC 6111  
| | | | |  
Qy 6121 TCTGGGTTTACACAGGCAAGCTACAGCCCGAAGGGGTGAGGCCCTCAGCCCTGTAG 6180  
| | | | |  
Db 6112 TCTGGGTTTACACAGGCAAGCTACAGCCCGAAGGGGTGAGGCCCTCAGCCCTGTAG 6171  
| | | | |  
Qy 6181 CTCAACCACTTGAACCCACGACAAAGGGGCTCCCAAGCACTGTGAAGCTTGACAAAG 6240  
| | | | |  
Db 6172 CTCAACCACTTGAACCCACGACAAAGGGGCTCCCAAGCACTGTGAAGCTTGACAAAG 6231  
| | | | |  
Qy 6241 CCACTGAGAGGGGAGCTGCGGCGCCAAAGCAGCAGCGCCCGTGAAGCTTGGCGGGAGGC 6300  
| | | | |  
Db 6232 CCACTGAGAGGGGAGCTGCGGCGCCAAAGCAGCAGCGCCCGTGAAGCTTGGCGGGAGGC 6291  
| | | | |  
Qy 6301 GCGCCACTTCCCAACACTGCGGCGCTGCTGAGAGCAGAGCCCTGTCAAGCCGCTCT 6360  
| | | | |  
Db 6292 GCGCCACTTCCCAACACTGCGGCGCTGCTGAGAGCAGAGCCCTGTCAAGCCGCTCT 6351  
| | | | |  
Qy 6361 CCAAGCGGCCCAAGGGTCAAAAGTCAACAGCGGGTGTCACTGTGCCAGCACTAG 6420  
| | | | |

Db 6352 CCAAGACCGCCGAGGGATCAAAAGGTCAACAAGCGGGTGTCAACCTTGCCCAAGACATCAAG 6411  
QY 6421 TGAAGTCATACACAGAGACTACACCCGGGACCAACCAAGCAGCTCAGAGCCACCTCGCC 6480  
Db 6412 TGAAGTCATACACAGAGACTACACCCGGGACCAACCAAGCAGCTCAGAGCCACCTCGCC 6471  
QY 6481 CCGCCCCCTCTACTCTCTCTCTGAGGACAGCTGCCCCCTGTGAGACTCCGCGCCACCC 6540  
Db 6472 CCGCCCCCTCTACTCTCTCTCTGAGGACAGCTGCCCCCTGTGAGACTCCGCGCCACCC 6531  
QY 6541 CAGTGACCTCTACCTCCGCCCCCGGACCAATGATGCCCGCCCGGTGACTCCCCCAAG 6600  
Db 6532 CAGTGACCTCTACCTCCGCCCCCGGACCAATGATGCCCGCCCGGTGACTCCCCCAAG 6591  
QY 6601 CCAAGAGGGGCAAGAGGTCTCAAGGCAAAACAAGAGTGGGTCTTGGGTGGTGGAGAG 6660  
Db 6592 CCAAGAGGGGCAAGAGGTCTCTCAAGGCAAAACAAGAGTGGGTCTTGGGTGGTGGAGAG 6651  
QY 6661 CCGTATTGAACTGTGTCCCAACCGAGGGCATGACGAGCCAGGGCACTCCGAGATGC 6720  
Db 6652 CCGTATTGAACTGTGTCCCAACCGAGGGCATGACGAGCCAGGGCACTCCGAGATGC 6711  
QY 6721 TGTGTACCCGCTGTGTACCGGGATGGGGAAACAAGAGCCAGAGGATGGGCTTCAAA 6771  
Db 6712 TGTGTACCCGCTGTGTACCGGGATGGGGAAACAAGAGCCAGAGGATGGGCTTCAAA 6771  
QY 6781 GTCTCCAGGCAACAAGCGCCGCAAGCTTCTTCAAGACCTGACCGAGAGCAATC 6840  
Db 6772 GTCTCCAGGCAACAAGCGCCGCAAGCTTCTTCAAGACCTGACCGAGAGCAATC 6831  
QY 6841 CCGCATGTCTCAAGTCCAGAGAGCAAGATCAACAAGAGCTGAAACCAACAACCGGAA 6900  
Db 6832 CCGCATGTCTCAAGTCCAGAGAGCAAGATCAACAAGAGCTGAAACCAACAACCGGAA 6891  
QY 6901 TGAAGCTGAATACATATACAGCAGCTTGGAGCGAGATCTTCAATATGCCCGCATAC 6960  
Db 6892 TGAAGCTGAATACATATACAGCAGCTTGGAGCGAGATCTTCAATATGCCCGCATAC 6951  
QY 6961 CGGAAAGGCGCTTATGACCTATAGAGCAAGCGGGGTGAGAAACATGCCAGCAACAT 7020  
Db 6952 CGGAAAGGCGCTTATGACCTATAGAGCAAGCGGGGTGAGAAACATGCCAGCAACAT 7011  
QY 7021 GGGGCTGAGAGCCATTAATTAAGAAAGCATCATGGGTAAATATGACAGTGGAGAGATC 7080  
Db 7012 GGGGCTGAGAGCCATTAATTAAGAAAGCATCATGGGTAAATATGACAGTGGAGAGATC 7071  
QY 7081 CCGCGCGCTCAGCGCCATGCTTTTAAACCTCTGAATGCCAGTGCAGCCTGCGCTGC 7140  
Db 7072 CCGCGCGCTCAGCGCCATGCTTTTAAACCTCTGAATGCCAGTGCAGCCTGCGCTGC 7131  
QY 7141 TATGCCCATTAACCGCTGTGACGAGAGAGTGAACAACATCACTTCCGCAAGTGGCCG 7200  
Db 7132 TATGCCCATTAACCGCTGTGACGAGAGAGTGAACAACATCACTTCCGCAAGTGGCCG 7191  
QY 7201 CCGGAAAGGCAAGGTCTTGGGAGAACCCAGCAGCCGAAAGCAATGCCCGGCCCGCGG 7260  
Db 7192 CCGGAAAGGCAAGGTCTTGGGAGAACCCAGCAGCCGAAAGCAATGCCCGGCCCGCGG 7251  
QY 7261 CCTGAGCATGTGGAGACCGGCAACCTCTGTCTTCAAGTGCATCTGAGAGAGAGACTGCAA 7320  
Db 7252 CCTGAGCATGTGGAGACCGGCAACCTCTGTCTTCAAGTGCATCTGAGAGAGAGACTGCAA 7311  
QY 7321 CCGCGGAGAGCGCGCTCAACAACCGCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 7380  
Db 7312 CCGCGGAGAGCGCGCTCAACAACCGCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 7371  
QY 7381 GGCATTCCCTCAACACCCCTGATCATGAGGCTGACGAGCGGTGATGATGATGATGATG 7440  
Db 7372 GGCATTCCCTCAACACCCCTGATCATGAGGCTGACGAGCGGTGATGATGATGATGATG 7431  
QY 7441 CCAACCGGAGCTTCCCGCGGAGAGCGGGCCCTCTGCTGAGCCCAACAGCTTGGAGAG 7500  
Db 7432 CCAACCGGAGCTTCCCGCGGAGAGCGGGCCCTCTGCTGAGCCCAACAGCTTGGAGAG 7491

QY 7501 GGAGCCCAAGCCACTGCTGTGTGAGATGAGAGACATCTCCGAGAGCGAGTGAAGTCAAG 7560  
Db 7492 GGAGCCCAAGCCACTGCTGTGTGAGATGAGAGACATCTCCGAGAGCGAGTGAAGTCAAG 7551  
QY 7561 AACAGGCGGGGGGGGGCGGGCGGTGERSPERLUCTCAGGTCTCCAGGAGCAAGAGAA 7620  
Db 7552 AACAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 7599  
QY 7621 CCGCCCTTGAGAGAGCGGGGCGGTGCGACTCCCCCAACCAAGAGAGAGAGAGAGAGAG 7680  
Db 7600 CCGCCCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7659  
QY 7681 CCGCTGAGCTCCATCATCTGTCCGTCCAGAGCCGGCATCTTGGCTGTCTAAAGCTTT 7740  
Db 7660 CCGCTGAGCTCCATCATCTGTCCGTCCAGAGCCGGCATCTTGGCTGTCTAAAGCTTT 7719  
QY 7741 AACTAAGACTCCCGCCCGGGCTGTGTGAGACTTACTCAGGGGATGTTTACTTG 7800  
Db 7720 AACTAAGACTCCCGCCCGGGCTGTGTGAGACTTACTCAGGGGATGTTTACTTG 7779  
QY 7801 GTGCTGGGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7860  
Db 7780 GTGCTGGGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7839  
QY 7861 ACAAGAGCGGCGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7920  
Db 7840 ACAGAGCGGCGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7899  
QY 7921 TGCCTCCCGGAAATGATTTGAAACCAAGTCTTAACTGAGCTCGAGCCCGGCGCTT 7980  
Db 7900 TGCCTCCCGGAAATGATTTGAAACCAAGTCTTAACTGAGCTCGAGCCCGGCGCTT 7959  
QY 7981 CCTCCGCTCCCATATCCGCTTATAGCCTCTGAGACATGAGAGAGAGAGAGAGAGAGAG 8040  
Db 7960 CCTCCGCTCCCATATCCGCTTATAGCCTCTGAGACATGAGAGAGAGAGAGAGAGAGAG 8019  
QY 8041 CCAAGTGCCTCGTCCGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8100  
Db 8020 CCAAGTGCCTCGTCCGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8079  
QY 8101 GTCAAGCCAGTGTGGGCGGACAAAGGGCCAGGTGCGGCTGGGGGAGAACGATGCTCCG 8160  
Db 8080 GTCAAGCCAGTGTGGGCGGACAAAGGGCCAGGTGCGGCTGGGGGAGAACGATGCTCCG 8139  
QY 8161 GGAATGGAATGTTTTTTTCAACATGTTGCGGAGCGGTGGAGAGAGAGAGAGAGAGAG 8220  
Db 8140 GGAATGGAATGTTTTTTTCAACATGTTGCGGAGCGGTGGAGAGAGAGAGAGAGAGAG 8199  
QY 8221 AATGATGTGTGGTTTACAGGGATATTTTGAATACCTTCAATGATTAATCAGATGTT 8280  
Db 8200 AATGATGTGTGGTTTACAGGGATATTTTGAATACCTTCAATGATTAATCAGATGTT 8259  
QY 8281 TTAAGCAAGAGAGACTTACCCAGATATTAATGCTGTGTGCTTTTGAATCTTGTCTTACG 8340  
Db 8260 TTAAGCAAGAGAGACTTACCCAGATATTAATGCTGTGTGCTTTTGAATCTTGTCTTACG 8319  
QY 8341 TTCAAGAGCGGTGTGAGAGCCGACAGTGGTGAACCTCATCACTCGAGAGACCAAGGGG 8400  
Db 8320 TTCAAGAGCGGTGTGAGAGCCGACAGTGGTGAACCTCATCACTCGAGAGACCAAGGGG 8379  
QY 8401 GGGGACTGCTCG-TCAAGCCCGCGGTGTGCTCCCTCCCTCCCTTCTTGGGAGAAATGA 8459  
Db 8380 GGGGACTGCTCGTCAAGCCCGCGGTGTGCTCCCTCCCTCCCTTCTTGGGAGAAATGA 8439  
QY 8460 ATTCAATGCGATTTCTGTGGCCGCAATTTGACAGAGGTGTGTATTTCTGTCAATTTAC 8519  
Db 8440 ATTCAATGCGATTTCTGTGGCCGCAATTTGACAGAGGTGTGTATTTCTGTCAATTTAC 8499  
QY 8520 ACGTGTCTTAATTAATAAGCAATTAATCTCA 8553  
Db 8500 ACGTGTCTTAATTAATAAGCAATTAATCTCA 8533

RESULT 7  
ACA62249 standard; cDNA; 8686 BP.  
ACA62249;  
12-AUG-2003 (first entry)  
cDNA encoding human nuclear receptor corepressor SMRTE.  
Human; ss; gene; SMRTE; nuclear receptor corepressor; gene therapy;  
tissue typing; cancer.  
Homo sapiens.  
Key Location/Qualifiers  
CDS 157..7680  
/tag= a  
/product= "SMRTE"  
US2003027137-A1.  
06-FEB-2003.  
27-MAR-2001; 2001US-00819104.  
29-MAR-2000; 2000US-0193138P.  
(CHEN/) CHEN J D.  
Chen JD;  
WPI; 2003-466139/44.  
P-PSDB; ABU61812.  
New SMRTE proteins and nucleic acids, useful in gene therapy, predictive  
medicine, therapeutic or prophylactic treatment, chromosome mapping,  
tissue typing and in forensic biology.  
Claim 2; Page 32-41; 90pp; English.  
The invention relates to an isolated SMRTE nucleic acid molecule. The  
nucleic acids are useful in gene therapy, as hybridisation probes for  
identifying SMRTE-encoding nucleic acid molecules and as primers for  
amplifying of SMRTE nucleic acid molecules. The polypeptides are useful  
as immunogens to raise anti-SMRTE antibodies. The SMRTE molecules are  
useful as targets for discovering and developing modulating agents to  
regulate a variety of cellular processes, in screening assays, in  
chemosensitive medicine, in therapeutic or prophylactic treatment, in  
chromosome mapping, tissue typing and in forensic identification of a  
biological sample. Modulators of SMRTE are useful for treating or  
preventing a condition associated with aberrant SMRTE protein or nucleic  
acid expression or activity, such as cancer. The present sequence  
represents cDNA encoding the human nuclear receptor corepressor SMRTE  
Sequence 8686 BP; 1881 A; 3069 C; 2562 G; 1174 T; 0 U; 0 Other;  
Query Match 97.2%; Score 8324.6; DB 8; Length 8686;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 8472; Conservative 1; Mismatches 34; Indels 88; Gaps 4;  
QY 1 CATGTCGGGCTTCACACAGCTTGTGGCAGACGCTGAGGGCCACTGAGCCCGCTACCC 60  
DB 156 CATGTCGGGCTTCACACAGCTTGTGGCAGACGCTGAGGGCCACTGAGCCCGCTACCC 215  
QY 61 GCCCAGACGCTTTCTTACCCAGTGCAGATCCCGGAGCGCAGACGAGAGTGGGCTCTT 120  
DB 216 GCCCAGACGCTTTCTTACCCAGTGCAGATCCCGGAGCGCAGACGAGTGGGCTCTT 275  
QY 121 GGAGTACAGACGACGCTCCCGAGATGAGCTCCAGACCTGTGGCGGGCTCATATCCA 180  
DB 276 GGAGTACAGACGACGCTCCCGAGATGAGCTCCAGACCTGTGGCGGGCTCATATCCA 335

QY 181 GCCCAGACGACGAGGCGCTTCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGCTCCCA 240  
DB 336 GCCCAGACGACGAGGCGCTTCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGCTCCCA 395  
QY 241 GGAGCTCAACCTGCGGGCCAGAGTCCCACTCATACCTGCGCCAGAGCTGGGGAAAGTCAGAGT 300  
DB 396 GGAGCTCAACCTGCGGGCCAGAGTCCCACTCATACCTGCGCCAGAGCTGGGGAAAGTCAGAGT 455  
QY 301 GGAGTTCAATTGAAGCAAGCGCCCTCGGCTAGAGCTGTGCTGTGACCCCTGCTCGAC 360  
DB 456 GGAGTTCAATTGAAGCAAGCGCCCTCGGCTAGAGCTGTGCTGTGACCCCTGCTCGAC 515  
QY 361 GTCAACCTGCTGTGACGAGGCGCAGGCTGTGCGGAACTTGAAGCCTTCAAGAACCTGAG 420  
DB 516 GTCAACCTGCTGTGACGAGGCGCAGGCTGTGCGGAACTTGAAGCCTTCAAGAACCTGAG 575  
QY 421 CTTGACGGGCAAGCTGGAAACCGGTGTCTCCCCCGACGCCCGGACACTGACCCCTGAGT 480  
DB 576 CTTGACGGGCAAGCTGGAAACCGGTGTCTCCCCCGACGCCCGGACACTGACCCCTGAGT 635  
QY 481 GGAGCTGTGCGGACCGGCTGTCCAGAGAGGCTGTACAGAACATGACCGGCTGGA 540  
DB 636 GGAGCTGTGCGGACCGGCTGTCCAGAGAGGCTGTACAGAACATGACCGGCTGGA 695  
QY 541 CCGAGAGATCAACCATGTGTAAGACAGACAGATCTTAACTGAAGAAAGACGAAACAGCT 600  
DB 696 CCGAGAGATCAACCATGTGTAAGACAGACAGATCTTAACTGAAGAAAGACGAAACAGCT 755  
QY 601 GGAGAGAGAGGCTGTGCAAGCCCGCCGACCTGAGAAAGCCGTTGATACCGCGCCATGGA 660  
DB 756 GGAGAGAGAGGCTGTGCAAGCCCGCCGACCTGAGAAAGCCGTTGATACCGCGCCATGGA 815  
QY 661 GTCGAAGCACCGCAGGCTGTGTGACAGATCATACGACGAAACCGGAAAGGCTGAAGC 720  
DB 816 GTCGAAGCACCGCAGGCTGTGTGACAGATCATACGACGAAACCGGAAAGGCTGAAGC 875  
QY 721 TGACATGTGATTTGGAAGGCTGTGGGCCCCAGGTGAGAGCTGCCGTATCAACACAGCC 780  
DB 876 TGACATGTGATTTGGAAGGCTGTGGGCCCCAGGTGAGAGCTGCCGTATCAACACAGCC 935  
QY 781 CTTCCGACACCGGGGAGTTCATGAGAACATCAAAATTAACGAGCGATGCGGAAAGAGT 840  
DB 936 CTTCCGACACCGGGGAGTTCATGAGAACATCAAAATTAACGAGCGATGCGGAAAGAGT 995  
QY 841 AATCTTGTACTTCAAGAGAGGAATACGCTCGAACAATGAAAGACGAAGTTCTGCA 900  
DB 996 AATCTTGTACTTCAAGAGAGGAATACGCTCGAACAATGAAAGACGAAGTTCTGCA 1055  
QY 901 GCGCTATGACCAAGCTCATGAGAGGCTTTGAAAAAAGGTGAGCGCATCGAAAAACACCC 960  
DB 1056 GCGCTATGACCAAGCTCATGAGAGGCTTTGAAAAAAGGTGAGCGCATCGAAAAACACCC 1115  
QY 961 GCGCGGGGGGCGAAGAGAGCAAGGTGCGGAGTACGAAAGCAAGTTCCCTGAGAT 1020  
DB 1116 GCGCGGGGGGCGAAGAGAGCAAGGTGCGGAGTACGAAAGCAAGTTCCCTGAGAT 1175  
QY 1021 CCGCAAGACGCGGAGCTGTGACGAGCGCATGCAAGAGGTTGGGCGACGCGGGCAGTGG 1080  
DB 1176 CCGCAAGACGCGGAGCTGTGACGAGCGCATGCAAGAGGTTGGGCGACGCGGGCAGTGG 1232  
QY 1081 GCTGTTCATGTGGCGCGCCGACGAGCAAGAGTGTCAAGATCATGATGGCTCTTC 1140  
DB 1233 GCTGTTCATGTGGCGCGCCGACGAGCAAGAGGTTCAAGATCATGATGGCTCTTC 1292  
QY 1141 AGAGAGAGAGAACTGTGAGAGAGAGATGCGCAGGTGCGCGGATCCGCGCCATGCTGA 1200  
DB 1293 AGAGAGAGAGAACTGTGAGAGAGAGATGCGCAGGTGCGCGGATCCGCGCCATGCTGA 1352  
QY 1201 GCAAGCTGACAGACGAGCATCAAGTTTATCAACAATGAAACGGGCTTATGCGGACCCCAT 1260  
DB 1353 GCAAGCTGACAGACGAGCATCAAGTTTATCAACAATGAAACGGGCTTATGCGGACCCCAT 1412



QY 1261 GAAGGTGTACAAAGACCGCCAGGTCAATGACATGTGAGTGAACAGAGAGAACTTT 1320  
DB 1413 GAAGGTGTACAAAGACCGCCAGGTCAATGACATGTGAGTGAACAGAGAGAACTTT 1472  
QY 1321 CCGGGGAAAGTTTCAATGACGATCCCAAGAACTTTGGCTGATGCAATCTTCTGGAGAG 1380  
DB 1473 CCGGGGAAAGTTTCAATGACGATCCCAAGAACTTTGGCTGATGCAATCTTCTGGAGAG 1532  
QY 1381 GAAGACATGGCTGATGGCTCTCTATTATCTAATCTGATCAAGAAAGAACTTATA 1440  
DB 1533 GAAGACATGGCTGATGGCTCTCTATTATCTAATCTGATCAAGAAAGAACTTATA 1592  
QY 1441 GAGCTGTGTGAGACGAGCTATTCGCGCGCGCAAGAGCAGCAGCAACACAGAGCA 1500  
DB 1593 GAGCTGTGTGAGACGAGCTATTCGCGCGCGCAAGAGCAGCAGCAACACAGAGCA 1652  
QY 1501 GACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1560  
DB 1653 GACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1712  
QY 1561 AGATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
DB 1713 AGATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1772  
QY 1621 CGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
DB 1773 CGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1832  
QY 1681 GAAAGAGCTGTGGCTCTCAAGAGCGCAAACTGCAACAGCAGAGAGAGAGAGAG 1740  
DB 1833 GAAAGAGCTGTGGCTCTCAAGAGCGCAAACTGCAACAGCAGAGAGAGAGAGAG 1892  
QY 1741 CGGCAATCAACCGCTCAATGGCTTATGAGGCGCAACAGCAGAGAGAGAGAGAG 1800  
DB 1893 CGGCAATCAACCGCTCAATGGCTTATGAGGCGCAACAGCAGAGAGAGAGAGAG 1952  
QY 1801 GAGCGCGAGCTGAGCTCTCAATGAGCTGATGAGAGTTCTCGCTGACAGAGAGAG 1860  
DB 1953 GAGCGCGAGCTGAGCTCTCAATGAGCTGATGAGAGTTCTCGCTGACAGAGAGAG 2012  
QY 1861 GAAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
DB 2013 GAAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2072  
QY 1921 GATGGGCTCAAGAGCTGATGAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980  
DB 2073 GATGGGCTCAAGAGCTGATGAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 2132  
QY 1981 GAACCTCGATGAGATCTTGACAGACAGCAAGCTGAGAGAGAGAGAGAGAGAGAG 2040  
DB 2133 GAACCTCGATGAGATCTTGACAGACAGCAAGCTGAGAGAGAGAGAGAGAGAGAG 2192  
QY 2041 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
DB 2193 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2252  
QY 2101 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160  
DB 2253 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2312  
QY 2161 GAGTGAAGCTTACATGCTCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220  
DB 2313 GAGTGAAGCTTACATGCTCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2321  
QY 2221 CACTGTCAACAAACAGTCAAGACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280  
DB 2322 CACTGTCAACAAACAGTCAAGACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2381  
QY 2281 GAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
DB 2382 GAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2441  
QY 2341 AGGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 2400

DB 2442 AGGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 2501  
QY 2401 TGAAGCCACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460  
DB 2502 TGAAGCCACCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2561  
QY 2461 TGTGTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
DB 2562 TGTGTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2621  
QY 2521 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
DB 2622 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2681  
QY 2581 GCTCCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
DB 2682 GCTCCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2741  
QY 2641 GAGAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
DB 2742 GAGAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2801  
QY 2701 CGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760  
DB 2802 CGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2861  
QY 2761 CTGCAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
DB 2862 CTGCAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2921  
QY 2821 AAGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
DB 2922 AAGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2981  
QY 2881 ACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
DB 2982 ACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3041  
QY 2941 AGTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000  
DB 3042 AGTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3101  
QY 3001 GCGACCGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
DB 3102 GCGACCGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3161  
QY 3061 GGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3100  
DB 3162 GGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3221  
QY 3101 ----GCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3156  
DB 3222 CCAGAGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3281  
QY 3157 CCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3216  
DB 3282 CCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3341  
QY 3217 CTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3276  
DB 3342 CTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3401  
QY 3277 CCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3336  
DB 3402 CCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3461  
QY 3337 CAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3396  
DB 3462 CAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3521  
QY 3397 GCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3456

Db 3522 GCTCCAGCTCCCGTACTCAGAGCATGCCAAAGGCCCCCGGTGGGCCTGTGCACCATGSGGCT 3581  
Qy 3457 GCCCTGCGCATGAGACCCCAAAAAGCTGGCACTTTCAAGCGGAGTGAAGCAGAGCAAGCT 3516  
Db 3582 GCCCTGCGCATGAGACCCCAAAAAGCTGGCACTTTCAAGCGGAGTGAAGCAGAGCAAGCT 3641  
Qy 3517 GTCCCAACGGGGCCAGGCTGGGCGCACCGGAGAGCTTGGGGGTGCCCAACAGCCAGAGAGC 3576  
Db 3642 GTCCCAACGGGGCCAGGCTGGGCGCACCGGAGAGCTTGGGGGTGCCCAACAGCCAGAGAGC 3701  
Qy 3577 GTCCGTGCTGAAGGGAGCAGAGCTTGGGCTCAAGTTCCGGGCGGGAACATCAACAAAGGAT 3636  
Db 3702 GTCCGTGCTGAAGGGAGCAGAGCTTGGGCTCAAGTTCCGGGCGGGAACATCAACAAAGGAT 3761  
Qy 3637 TCCCAAGCAGCGGGTGCCTTCGAGACGCGCATCAATACCGGCTCATCAACCAAGC 3696  
Db 3762 TCCCAAGCAGCGGGTGCCTTCGAGACGCGCATCAATACCGGCTCATCAACCAAGC 3821  
Qy 3697 CACGCGACGTGACGTCTGTATCAGAGGGCAACATCAACAGAGATCATCGGCGAGAGACGCC 3756  
Db 3822 CACGCGACGTGACGTCTGTATCAGAGGGCAACATCAACAGAGATCATCGGCGAGAGACGCC 3881  
Qy 3757 GAGTGCCTTGAACCGGCGCGGAGAGACAGCGCTGCGCAAGGGGCAAGTATTAAGAAAG 3816  
Db 3882 GAGTGCCTTGAACCGGCGCGGAGAGACAGCGCTGCGCAAGGGGCAAGTATTAAGAAAG 3941  
Qy 3817 CAAGAAAGGCGCAAGTCTTCTCTATAGAGGTGGACATGTCTGTAGCCAGTGTCTCAAGGA 3876  
Db 3942 CAAGAAAGGCGCAAGTCTTCTCTATAGAGGTGGACATGTCTGTAGCCAGTGTCTCAAGGA 4001  
Qy 3877 GAGCGGCAAGACAGCTCAGAGACCCCGCCATGAAACGGCGCGCCCAAGCGCATTAAGA 3936  
Db 4002 GAGCGGCAAGACAGCTCAGAGACCCCGCCATGAAACGGCGCGCCCAAGCGCATTAAGA 4061  
Qy 3937 CATGATGAAGGGCGGCTGGGAGAGCCATCTCTCAAGCAGCATGAAAGGTCTCATGGG 3996  
Db 4062 CATGATGAAGGGCGGCTGGGAGAGCCATCTCTCAAGCAGCATGAAAGGTCTCATGGG 4121  
Qy 3997 CCGTGCATTCGCGCGCGAGCGCACAGCGCCCAACCACTCAAAAGAGCAGCACATCG 4056  
Db 4122 CCGTGCATTCGCGCGCGAGCGCACAGCGCCCAACCACTCAAAAGAGCAGCACATCG 4181  
Qy 4057 CCGGTCTCATCACAAGAGATCTTGGTCTTACGTGAGGGCACAAGAGAACTACTTGG 4116  
Db 4182 CCGGTCTCATCACAAGAGATCTTGGTCTTACGTGAGGGCACAAGAGAACTACTTGG 4241  
Qy 4117 TCGGGAGGCGAAGCTCTTAAAGGGGAGGGGCAAGCTCCGCGCCCAACCGCCCTCAAGGGA 4176  
Db 4242 TCGGGAGGCGAAGCTCTTAAAGGGGAGGGGCAAGCTCCGCGCCCAACCGCCCTCAAGGGA 4301  
Qy 4177 CCTGACCGAGGCTTCAAGACGCGAGGCTTGGGCCCTTGAAGCTGAAGCGGGCCATGA 4236  
Db 4302 CCTGACCGAGGCTTCAAGAGCGCAAGCGCTTGGGCCCTTGAAGCTGAAGCGGGCCATGA 4361  
Qy 4237 GGGCTGTGGTCAAGTGAAGAGGGCGGGCGCTTCATCCATGAGATCCGCGAGGA 4296  
Db 4362 GGGCTGTGGTCAAGTGAAGAGGGCGGGCGCTTCATCCATGAGATCCGCGAGGA 4421  
Qy 4297 GCTGCGGACAGCGCGAGCTGCGCTGGCGCGGGCGCGCTCAAGAGGGGCTCCATAC 4356  
Db 4422 GCTGCGGACAGCGCGAGCTGCGCTGGCGCGGGCGCGCTCAAGAGGGGCTCCATAC 4481  
Qy 4357 GCAGGCGACCCCGCTCAAGTACGACCGGCGCGTTCACACTGGCTCCAAAGACACA 4416  
Db 4482 GCAGGCGACCCCGCTCAAGTACGACCGGCGCGTTCACACTGGCTCCAAAGACACA 4541  
Qy 4417 CGTACGCTTCCTCATCGGAGCGCGCGCGGAGCTTCCCAACCGTGCACCGCGTGAATG 4476  
Db 4542 CGTACGCTTCCTCATCGGAGCGCGCGCGGAGCTTCCCAACCGTGCACCGCGTGAATG 4601  
Qy 4477 GATGGCGAAGCGCGGGGCACTGGAAAGTGGCTCTACAGAGAGAGCACTGAAGAGCGGCG 4536  
Db 4602 GATGGCGAAGCGCGGGGCACTGGAAAGTGGCTCTACAGAGAGAGCACTGAAGAGCGGCG 4661

Qy 4537 AGGAGCCGCGACAGAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCGATCATTTGTCTGA 4596  
Db 4662 AGGAGCCGCGACAGAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCGATCATTTGTCTGA 4721  
Qy 4537 GCTGGGTAAAGCGCGGCGAGAGCCCTCTGAACCTATGAGAACACAGGGGACCTTTGGCGG 4656  
Db 4722 GCTGGGTAAAGCGCGGCGAGAGCCCTCTGAACCTATGAGAACACAGGGGACCTTTGGCGG 4781  
Qy 4657 CCACCTCCACAGAGGTGGCCCGTGAACCATGCGGGAGGCCACGCGCGCTTCAGAGAGG 4716  
Db 4782 CCACCTCCACAGAGGTGGCCCGTGAACCATGCGGGAGGCCACGCGCGCTTCAGAGAGG 4841  
Qy 4717 CAGCCTTTGCTTCAGCAAGGCAATCCCAAGACGAAAGGTGACGTGAGCGCTGTGAGAT 4776  
Db 4842 CAGCCTTTGCTTCAGCAAGGCAATCCCAAGACGAAAGGTGACGTGAGCGCTGTGAGAT 4901  
Qy 4777 CGCCAAGTCCCGCGCACAGACCGGTGCCGAGACCAACCAACCCCACTTGGCCCTTAAGA 4836  
Db 4902 CGCCAAGTCCCGCGCACAGACCGGTGCCGAGACCAACCAACCCCACTTGGCCCTTAAGA 4961  
Qy 4837 GCACTGTCTCGGGCGTGAAGTGGGTGAACCTGTATGCAAGCCATTCGCCCTGAGCTT 4896  
Db 4962 GCACTGTCTCGGGCGTGAAGTGGGTGAACCTGTATGCAAGCCATTCGCCCTGAGCTT 5021  
Qy 4897 GCACCCGCACTTCATACCCCGGCGATCCCTGTGAGCGAGCGCTGCTTACTACTTGGCC 4956  
Db 5022 GCACCCGCACTTCATACCCCGGCGATCCCTGTGAGCGAGCGCTGCTTACTACTTGGCC 5081  
Qy 4957 CCGACACTGTGCGCCCAACCCCACTTACCGGCACTTGTACCAACCTTACCTCAATCCGGG 5016  
Db 5082 CCGACACTGTGCGCCCAACCCCACTTACCGGCACTTGTACCAACCTTACCTCAATCCGGG 5141  
Qy 5017 CTACCCCGACAGCGGGCGCTGAGAGAACGGGACAGACATCAATGACTACATCACTTC 5076  
Db 5142 CTACCCCGACAGCGGGCGCTGAGAGAACGGGACAGACATCAATGACTACATCACTTC 5201  
Qy 5077 GCAGAGATGACCAACAACAGGCCACCGGCATGCGGAGGCTGATATGCTGAGGGG 5136  
Db 5202 GCAGAGATGACCAACAACAGGCCACCGGCATGCGGAGGCTGATATGCTGAGGGG 5261  
Qy 5137 CCTTCGCGCCCGGAGTCTGTGGTCACTCAACTACGCTGGGTCCCGGAGGATCAT 5196  
Db 5262 CCTTCGCGCCCGGAGTCTGTGGTCACTCAACTACGCTGGGTCCCGGAGGATCAT 5321  
Qy 5197 GCACTGTCCCAAGGCGCACACTGCTGTGCTGTGCGCCCGACACGAGGACCCAGC 5256  
Db 5322 GCACTGTCCCAAGTGCACACTGCTGTGCTGTGCGCCCGACACGAGGACCCAGC 5381  
Qy 5257 CACGCGCATGAGCGGCTTGTCTTACCTTCCCAACGCGGCGCCAGCCCTTCAAGAGCGGCA 5316  
Db 5382 CACGCGCATGAGCGGCTTGTCTTACCTTCCCAACGCGGCGCCAGCCCTTCAAGAGCGGCA 5441  
Qy 5317 CAGAGCTTCCCACTTCTCCCGAGAGGTCCAAACAACCTTGAACAAAACCAACACAGCTC 5376  
Db 5442 CAGAGCTTCCCACTTCTCCCGAGAGGTCCAAACAACCTTGAACAAAACCAACACAGCTC 5501  
Qy 5377 CTCCGTCCGAGCGGAGCGAGACCGGGATTCGAGCGGGAGCCGGGATCGGAGCGGGAAAA 5436  
Db 5502 CTCCGTCCGAGCGGAGCGAGACCGGGATTCGAGCGGGAGCCGGGATTCGGAGCGGGAAAA 5561  
Qy 5437 GTCCATCTCTCAAGTTCACACGACGAGTGAAGACGCACTTGTGAGACCTTGTATCAGA 5496  
Db 5562 GTCCATCTCTCAAGTTCACACGAGCGGTGAAGACGCACTTGTGAGACCTTGTATCAGA 5621  
Qy 5497 GCAGAGAGCGGCGACAGCGGCGAGAGCGGGGGTGGGGGACAGCAAGCGCGCGCGCG 5556  
Db 5622 GCAGAGAGCGGCGACAGCGGCGAGAGCGGGGGTGGGGGACAGCAAGCGCGCGCGCG 5681  
Qy 5557 CTCCCACTCCCATGCGCACAGCACTCGCCCATCTCCCTCGGACCCAGAGATGCCCTCA 5616  
Db 5682 CTCCCACTCCCATGCGCACAGCACTCGCCCATCTCCCTCGGACCCAGAGATGCCCTCA 5741

QY 5617 GCAGAGACCCAGTGTGCTTCAACAAGGCAATGAAGGATATCATGCCGTGTGAGGCC 5676  
DB 5742 GCAGAGACCCAGTGTGCTTCAACAAGGCAATGAAGGATATCATGCCGTGTGAGGCC 5801  
QY 5677 CAGCAAGCCCAAGGTCTGTAGGTTCACCTTCACCTCTCAACCCGTTGCCAGCTGCAC 5736  
DB 5802 CAGCAAGCCCAAGGTCTGTAGGTTCACCTTCACCTCTCAACCCGTTGCCAGCTGCAC 5861  
QY 5737 ATTCCCAACCTGACCACTGACCACTGAGGAGGCACTTGATGAGGATCTAACCTACCT 5796  
DB 5862 ATTCCCAACCTGACCACTGACCACTGAGGAGGCACTTGATGAGGATCTAACCTACCT 5921  
QY 5797 CATGAGGCCGTGTCTGTGCTGACCAAGAGAGGCCCGAGTGCACGAGCCAGAGGCCCGC 5856  
DB 5922 CATGAGGCCGTGTCTGTGCTGACCAAGAGAGGCCCGAGTGCACGAGGCCCGC 5981  
QY 5857 AGCAGACACCGGCAATGCTTCTGCGCAAGCCCCAGCCGCTCGAGGTGAGGCCGC 5916  
DB 5982 AGCAGACACCGGCAATGCTTCTGCGCAAGCCCCAGCCGCTCGAGGTGAGGCCGC 6041  
QY 5917 CTCTCTCCCAAGAGAGGCTCGAGGCCCGAGCCCTTAAGTGCCTCTGCTGTGAGCCAGC 5976  
DB 6042 CTCTCTCCCAAGAGAGGCTCGAGGCCCGAGCCCTTAAGTGCCTCTGCTGTGAGCCAGC 6101  
QY 5977 CACCAATGCCCCGACACCTGTGCAAGAACTCGCACTCAACGCGAGCCGAGCCGAGC 6036  
DB 6102 CACCAATGCCCCGACACCTGTGCAAGAACTCGCACTCAACGCGAGCCGAGCCGAGC 6161  
QY 6037 GGGGCCACTGCTCGGCTCGGACCCGCAACGGGAAAGACTCAAGTAACCTTTTC 6096  
DB 6162 GGGGCCACTGCTCGGCTCGGACCCGCAACGGGAAAGACTCAAGTAACCTTTTC 6221  
QY 6097 CATCCAGAACTGGAATCTCCGTTCTGTGGTTACCAAGGCAAGAGTTACAGCCCCGAGG 6156  
DB 6222 CATCCAGAACTGGAATCTCCGTTCTGTGGTTACCAAGGCAAGAGTTACAGCCCCGAGG 6281  
QY 6157 GGTGAGGCCGTGTGAGCTCAACCAAGTCTGACCCAGCAAGAGGAGCTCCCA 6216  
DB 6282 GGTGAGGCCGTGTGAGCTCAACCAAGTCTGACCCAGCAAGAGGAGCTCCCA 6341  
QY 6217 GCACTGGAAGAGCTGACAAAGAGCACTTGAAGGAGAGCTGAGCCCAAGAGCCAG 6276  
DB 6342 GCACTGGAAGAGCTGACAAAGAGCACTTGAAGGAGAGCTGAGCCCAAGAGCCAG 6401  
QY 6277 CCCCCGGAAGCTTGGGGGAGGAGCCGACCTTCCCAACATGAGCCGCTGCTGATGAG 6336  
DB 6402 CCCCCGGAAGCTTGGGGGAGGAGCCGACCTTCCCAACATGAGCCGCTGCTGATGAG 6461  
QY 6337 CCAGCCCTGTCAAGCCGCTGCTCAAGCCGACCGAGGAGTCAACGAGGAGT 6396  
DB 6462 CCAGCCCTGTCAAGCCGCTGCTCAAGCCGACCGAGGAGTCAACGAGGAGT 6521  
QY 6397 GGTCAACCTGAGCCAGACATCATGAGGATCATCAACAGAGTACACCCGAGCAACCC 6456  
DB 6522 GGTCAACCTGAGCCAGACATCATGAGGATCATCAACAGAGTACACCCGAGCAACCC 6581  
QY 6457 AAGAGAGCTCAAGGACACCTCTGCCCCCTCTTACTCTTCTCTGAGGAGCAGCTGCC 6516  
DB 6582 AAGAGAGCTCAAGGACACCTCTGCCCCCTCTTACTCTTCTCTGAGGAGCAGCTGCC 6641  
QY 6517 GGTCTGAGACTCGGCGGACCAACCAAGTACCTTACCTCCGAGCCCGAGACATGATG 6576  
DB 6642 GGTCTGAGACTCGGCGGACCAACCAAGTACCTTACCTCCGAGCCCGAGACATGATG 6701  
QY 6577 CCGGAGCCGTGAGCTCCCAACAGCGAAGGAGCAAGAGTCTCAAGAGCAAAAGAG 6636  
DB 6702 CCGGAGCCGTGAGCTCCCAACAGCGAAGGAGCAAGAGTCTCAAGAGCAAAAGAG 6761  
QY 6637 GTTCGCTTGTGGGTGTGAGGAGCGGTATTGAACCTGTGTCCCAACGAGAGGAGTGA 6696  
DB 6762 GTTCGCTTGTGGGTGTGAGGAGCGGTATTGAACCTGTGTCCCAACGAGAGGAGTGA 6821  
QY 6697 GAGAGCGAGGACTCCGAGAGTGTGTATCCGCTGTATCCGAGATGAGGAGACAGAC 6756

DB 6822 GAGAGCGAGGACTCCGAGAGTGTGTATCCGCTGTGTATCCGAGATGAGGAGACAGAC 6881  
QY 6757 GAGAGCCAGAGATGAGGCTCCAAAGTCTCAAGGCAACACAGGACGCGCAACCTTCT 6816  
DB 6882 GAGAGCCAGAGATGAGGCTCCAAAGTCTCAAGGCAACACAGGACGCGCAACCTTCT 6941  
QY 6817 CAGCAAGCTGACCGAGAGCAATCCGCAATGTCAAGTCCAAAGTCCAAAGAGATCAACA 6876  
DB 6942 CAGCAAGCTGACCGAGAGCAATCCGCAATGTCAAGTCCAAAGTCCAAAGAGATCAACA 7001  
QY 6877 GAAAGTGAACCCCAACCGGAATGAGCTGAAATCAATATCAGCCAGCTGAGAGG 6936  
DB 7002 GAAAGTGAACCCCAACCGGAATGAGCTGAAATCAATATCAGCCAGCTGAGAGG 7061  
QY 6937 GATCTTCAATATGCCCCGATCAACGGAACAGGCTTATGACCTTATGAAGCCAGCGGT 6996  
DB 7062 GATCTTCAATATGCCCCGATCAACGGAACAGGCTTATGACCTTATGAAGCCAGCGGT 7121  
QY 6997 GCAGAAATGACCAAGCAACATGAGGCTGAGAGCCATTAATTAAGAGCACTCATGG 7056  
DB 7122 GCAGAAATGACCAAGCAACATGAGGCTGAGAGCCATTAATTAAGAGCACTCATGG 7181  
QY 7057 TAAATATGACCAATGAGAAAGTCCCGCGCTCAGCGCCAAATGCTTTAACTTGTAA 7116  
DB 7182 TAAATATGACCAATGAGAAAGTCCCGCGCTCAGCGCCAAATGCTTTAACTTGTAA 7241  
QY 7117 TGCAGTGCAGGCTGCCCCGCTGTATGCCCCATTAACCGTGTGAGAGAGAGTACCA 7176  
DB 7242 TGCAGTGCAGGCTGCCCCGCTGTATGCCCCATTAACCGTGTGAGAGAGAGTACCA 7301  
QY 7177 CACACTCACCTGSCAAGTGGAGGAGGAGAAAGCAAGGTCTGTGGAGACCAGAGCCG 7236  
DB 7302 CACACTCACCTGSCAAGTGGAGGAGGAGAAAGCAAGGTCTGTGGAGACCAGAGCCG 7361  
QY 7237 AAAAGCCAAGTCCCGGCCCCGAGCTGTGATCTGAGGAGACCGGCAACCTTGTCTCT 7296  
DB 7362 AAAAGCCAAGTCCCGGCCCCGAGCTGTGATCTGAGGAGACCGGCAACCTTGTCTCT 7421  
QY 7297 AGTGCACTGAGAGGAGATGCAACCGCGGAGCGGCTCAACACCGGTGTGGAGGA 7356  
DB 7422 AGTGCACTGAGAGGAGATGCAACCGCGGAGCGGCTCAACACCGGTGTGGAGGA 7481  
QY 7357 CAGGCCCTGTCCGAGAGTTCACGCAATCCCTTAACCCCTGATCATGAGGCTGCA 7416  
DB 7482 CAGGCCCTGTCCGAGAGTTCACGCAATCCCTTAACCCCTGATCATGAGGCTGCA 7541  
QY 7417 GGGGAGTGTATGAGCTTCCCAACCCCAACGGGAGCTCCCGGAGGAGCGGCTTCG 7476  
DB 7542 GGGGAGTGTATGAGCTTCCCAACCCCAACGGGAGCTCCCGGAGGAGCGGCTTCG 7601  
QY 7477 TGGCCCCCAACAGCCTGAGGAGAGAGGCTCAAGCCATGCTCTGCTCGAGATACAGAC 7536  
DB 7602 TGGCCCCCAACAGCCTGAGGAGAGAGGCTCAAGCCATGCTCTGCTCGAGATACAGAC 7661  
QY 7537 ACTCTCCGACAGAGTGACTGAACAAGGAGGAGGAGGAGGAGTGTGAGTGTG 7596  
DB 7662 ACTCTCCGACAGAGTGACTGAACAAGGAGGAGGAGGAGGAGTGTGAGTGTG 7711  
QY 7597 TCAAGTCCAGAGCAACAGGAACGACCTGCAAGAGCGGAGCGCTGCACTCCCG 7656  
DB 7712 TCAAGTCCAGAGCAACAGGAACGACCTGCAAGAGCGGAGCGCTGCACTCCCG 7771  
QY 7657 AACCAAGAGAGAGCCCTGAGTCCGCTGCGCTCATCATCTGTCTGCTCAAGAGCG 7716  
DB 7772 AACCAAGAGAGAGCCCTGAGTCCGCTGCGCTCATCATCTGTCTGCTCAAGAGCG 7831  
QY 7717 GATCTTGTGCTGTCTTAAGCCCTTAACCTAAGCTCCCGCGGAGTGTGCTGTGAG 7776  
DB 7832 GATCTTGTGCTGTCTTAAGCCCTTAACCTAAGCTCCCGCGGAGTGTGCTGTGAG 7891  
QY 7777 CCTTACTAGGAGATGTTTAACTGTGTCTGAGGAAGGAGGAGGAGCGGAGGAG 7836

```
Db 7892 CCTACTCAGGGGATGTTTACTGCTGCTCGGGAAGGAGGGAGCGGAGGAG 7951
Qy 7837 GCACGGCAGGCGCTGTGGAGAGCCACACAGCGCGCCAGGGCGGCCCAAGAG 7896
Db 7952 GCACGGCAGGCGGTGTGGAGAGCCACACAGCGCGGCCAGGGCGGCCCAAGAG 8011
Qy 7897 GATGACGACGACCTTCACGCGCATGCTCTCCCGAATGCAATTGGAACCAAGTCTAA 7956
Db 8012 GATGACGACGACCTTCACGCGCATGCTCTCCCGAATGCAATTGGAACCAAGTCTAA 8071
Qy 7957 CTGAGCTGCGAGCGCGCGCGCTCTCCCTCCGCTCCCAATCCGCTTAAGCGCTGAGACG 8016
Db 8072 CTGAGCTGCGAGCGCGCGCGCTCTCCCTCCGCTCCCAATCCGCTTAAGCGCTGAGACG 8131
Qy 8017 ATGACGACGAGCGCTGTCCAGCGCCCAAGTGCCTGTTCCGGTCCCAAGACTGCCCA 8076
Db 8132 ATGACGACGAGCGCTGTCCAGCGCCCAAGTGCCTGTTCCGGTCCCAAGACTGCCCA 8191
Qy 8077 GCCAAGGATTTGCTGGAACCAAGTCAAGGCGAGGTGGCGGACAAAGGGCCAGGTGC 8136
Db 8192 GCCAAGGATTTGCTGGAACCAAGTCAAGGCGAGGTGGCGGACAAAGGGCCAGGTGC 8251
Qy 8137 GCTGGGGGGGAAAGGATGCTCCGAGACCTGACCTGTTTTTTTCAACATCCGTTGGCGAG 8196
Db 8252 GCTGGGGGGGAAAGGATGCTCCGAGACCTGACCTGTTTTTTTCAACATCCGTTGGCGAG 8311
Qy 8197 CGGTGGGAAAGGAAAGGAGATGAATGATGTGTGTTTACAGGGATATATTTTGTATAC 8256
Db 8312 CGGTGGGAAAGGAAAGGAGATGAATGATGTGTGTTTACAGGGATATATTTTGTATAC 8371
Qy 8257 CTTCAATGATTAATTCAAGATGTTTTTACGGAAGAAAGACCTTACCGATATTACTGCTGC 8316
Db 8372 CTTCAATGATTAATTCAAGATGTTTTTACGGAAGAAAGACCTTACCGATATTACTGCTGC 8431
Qy 8317 TGTGCTTTTGAATCTGCTGCTTACGTTTCAAGAGCGGTGGCGAGCACTGCGTGAAGCC 8376
Db 8432 TGTGCTTTTGAATCTGCTGCTTACGTTTCAAGAGCGGTGGCGAGCACTGCGTGAAGCC 8491
Qy 8377 CATCACTGCGAGGACCAAGGGGGCGGAGCTGCTGTCACGCGCGCTGTGCTCTCCCTC 8436
Db 8492 CATCACTGCGAGGACCAAGGGGGCGGAGCTGCTGTCACGCGCGCTGTGCTCTCCCTC 8551
Qy 8437 CCTCCCTTCCTTGGGCAAGATGAATTCAGTGCATATTCGTGGCGCGCATTTGGCGAGGG 8496
Db 8552 CCTCCCTTCCTTGGGCAAGATGAATTCAGTGCATATTCGTGGCGCGCATTTGGCGAGGG 8611
Qy 8497 TGTGCTGATTTCTGTCAATTACACAGTGTGTTTAAAGCGAATTAATCTCCAAA 8556
Db 8612 TGTGCTGATTTCTGTCAATTACACAGTGTGTTTAAAGCGAATTAATCTCCAAA 8671
Qy 8557 AAAAAAAAAAAAAA 8571
Db 8672 AAAAAAAAAAAAAA 8686

RESULT 8
ADL13811
ID ADL13811 standard; DNA; 8686 BP.
XX
AC ADL13811;
XX
DE 06-MAY-2004 (first entry)
XX
DE Osteoarthritis-associated polymorphic nucleotide #343.
XX
KW ds; gene; osteopathic; antiinflammatory; antiarthritis; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN W02003054166-A2.
XX
```

```
PD 03-JUL-2003.
XX
XX 19-DEC-2002; 2002WO-US041225.
XX
XX 20-DEC-2001; 2001US-0342603P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Jones KA, Schafer A;
XX
XX MPI; 2003-559141/52.
XX
XX Determining susceptibility of an individual to joint space narrowing,
XX osteophyte development and/or joint pain comprises identifying whether
XX the individual has at least one polymorphism in a polynucleotide encoding
XX a protein.
XX
XX Disclosure; SEQ ID NO 343; 297bp; English.
XX
XX The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polynucleotide encoding at least one of the protein
XX listed in the specification. The methods, composition and agent are
XX useful for modulating the susceptibility of an individual to joint space
XX narrowing and/or osteophyte development and/or joint pain that is
XX associated with a disease, preferably osteoarthritis. The cell line and
XX the non-human animal are useful for screening for an agent for diagnosing
XX an individual having susceptibility to joint space narrowing and/or
XX osteophyte development and/or joint pain. This sequence corresponds to
XX the polynucleotide encoding a protein listed in the specification. (Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pcr_sequences).
XX
XX Sequence 8686 BP; 1881 A; 3069 C; 2562 G; 1174 T; 0 U; 0 Other;
XX

Query Match 97.2%; Score 8324.6; DB 10; Length 8686;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 8472; Conservative 1; Mismatches 34; Indels 88; Gaps 4;

Qy 1 CATGTGGGCTTCCACAGCTTTGTGGACACAGCTGGAGGGCACTGAGCCCGCTACCC 60
Db 156 CATGTGGGCTTCCACAGCTTTGTGGACACAGCTGGAGGGCACTGAGCCCGCTACCC 215
Qy 61 GCCCAGACGCTTTCTTACCCAGTCCAGATGCGCCGAGACACACAGAGCTGCGGCTCT 120
Db 216 GCCCAGACGCTTTCTTACCCAGTCCAGATGCGCCGAGACACAGAGCTGCGGCTCT 275
Qy 121 GGAGTACGAGACCACTCCGCGACTATGCTCCGACTGTGCGCGGCTTCATCATCA 180
Db 276 GGAGTACGAGACCACTCCGCGACTATGCTCCGACTGTGCGCGGCTTCATCATCA 335
Qy 181 GCCCAGGCGGGAGAGCCCTCTGCTGTGATGTTCCAGCCCGGGAATGAACGTTCCA 240
Db 276 GGAGTACGAGACCACTCCGCGACTATGCTCCGACTGTGCGCGGCTTCATCATCA 335
Qy 336 GCCCAGGCGGGAGAGCCCTCTGCTGTGATGTTCCAGCCCGGGAATGAACGTTCCA 395
Db 241 GGAGTCCACCTGCGGGACAGTCCCACTCATCTGCGCCGAGCTGGGGAATGAGAGAT 300
Qy 336 GGAGTCCACCTGCGGGACAGTCCCACTCATCTGCGCCGAGCTGGGGAATGAGAGAT 455
Db 301 GGAGTTCATTGAAGCAAGCGCCCTCGGCTAGAGCTGTGCTGACCCCTGCTGCAAC 360
Qy 456 GGAGTTCATTGAAGCAAGCGCCCTCGGCTAGAGCTGTGCTGACCCCTGCTGCAAC 515
Db 361 GTACACCTCTGCTGCGACAGGCGCAAGCTGCGGAGTCTGAAGACTCACACAGAGACCTGAG 420
Qy 516 GTACACCTCTGCTGCGACAGGCGCGAGCTGCGGAGTCTGAAGACTCACACAGAGACCTGAG 575
Db 421 CCGTACGGGCAAGCTGGAACGGGTGTCTCCCGCCAGCCCGCGACACTGACCTGAGCT 480
Qy 576 CCGTACGGGCAAGCTGGAACGGGTGTCTCCCGCCAGCCCGCGACACTGACCTGAGCT 635
Db
```

481 GGAGCTGTGTCGCCCAACGGCTGTCCAAGAGAGCTGATCCAGAAATGGAACCGCGTGA 540  
636 GGAGCTGTGTGCGCCACCGCTGTCCAAGAGAGAGCTGATCCAGAAATGGAACCGCGTGA 695  
541 CCGAGAGATCAACATGTGTAGAGAGAGATCTTTAAAGCTGAGAGAGAGAGAGAGAGCT 600  
696 CCGAGAGATCAACATGTGTAGAGAGAGAGATCTTTAAAGCTGAGAGAGAGAGAGAGCT 755  
601 GGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
756 GGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 815  
661 GTGGAAGACACCGAGAGCTGTGTGAGATCATCTACGAGAGAGAGAGAGAGAGAGAGCT 720  
816 GTGGAAGACACCGAGAGCTGTGTGAGATCATCTACGAGAGAGAGAGAGAGAGAGAGCT 875  
721 TGCACATCGGATTTCTGAAAGGCTGTGGGGCCCAAGTGAAGCTGCCGTGTACACAGCC 780  
876 TGCACATCGGATTTCTGAAAGGCTGTGGGGCCCAAGTGAAGCTGCCGTGTACACAGCC 935  
781 CTCGCAACACCCGAGATATCAAGAGAAATCAAAATTAACCAAGAGAGAGAGAGAGCT 840  
936 CTCGCAACACCCGAGATATCAAGAGAAATCAAAATTAACCAAGAGAGAGAGAGCT 995  
841 AATCTTGATCTTCAAGAGAGAGAGATCAAGCTCGAGAAACAAATGGAAGAGAGAGAGCT 900  
996 AATCTTGATCTTCAAGAGAGAGAGATCAAGCTCGAGAAACAAATGGAAGAGAGAGAGCT 1055  
901 GCGCTATGACCAAGCTCATGAGAGGCTTTGAAAAAGGTGAGAGAGAGAGAGAGAGCT 960  
1056 GCGCTATGACCAAGCTCATGAGAGGCTTTGAAAAAGGTGAGAGAGAGAGAGAGAGCT 1115  
961 GCGCCGCGCGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
1116 CCGGCGCGCGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1175  
1021 CCGCAAGACAGCGGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
1176 CCGCAAGACAGCGGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1232  
1081 GCTGTTCATGTGCGGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140  
1233 GCTGTTCATGTGCGGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1292  
1141 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
1293 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1352  
1201 CCAAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
1353 CCAAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1412  
1261 GAAAGTGTACAAAGACCGGAGAGTGTACATGTGAGTGAAGAGAGAGAGAGAGAGCTT 1320  
1413 GAAAGTGTACAAAGACCGGAGAGTGTACATGTGAGTGAAGAGAGAGAGAGAGAGCTT 1472  
1321 CCGGAGAGAGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
1473 CCGGAGAGAGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1532  
1381 GAAAGACAGTGTGAGTGTGAGTGTCTTATTACTGAGTGAAGAGAGAGAGAGAGAGCTTAA 1440  
1533 GAAAGACAGTGTGAGTGTGAGTGTCTTATTACTGAGTGAAGAGAGAGAGAGAGAGCTTAA 1592  
1441 GAGCTGTGTGAGAGAGAGAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
1593 GAGCTGTGTGAGAGAGAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1652  
1501 GCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
1653 GCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1712  
1561 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620

1713 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1772  
1621 CGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
1773 CGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1832  
1681 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
1833 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1892  
1741 CCGCATCACCCGCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800  
1893 CCGCATCACCCGCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1952  
1801 GAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860  
1953 GAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2012  
1861 GGAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
2013 GGAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2072  
1921 GGTGGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980  
2073 GGTGGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2132  
1981 GAACTGTGATGAGATTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040  
2133 GAACTGTGATGAGATTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2192  
2041 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
2193 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2252  
2101 GAGAGATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160  
2253 GAGAGATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2312  
2161 GGTGAGAGCTTATCATGCTCTGAGAGATGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220  
2313 GGTGAGAGCTTATCATGCTCTGAGAGATGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2321  
2221 CACTGTCAACACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280  
2322 CACTGTCAACACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2381  
2281 GGAACAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
2382 GGAACAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2441  
2341 AGGCCCCAACCCCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400  
2442 AGGCCCCAACCCCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2501  
2401 TGAAGCCACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460  
2502 TGAAGCCACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2561  
2461 TGTGTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
2562 TGTGTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2621  
2521 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
2622 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2681  
2581 GCCCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
2682 GCCCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2741  
2641 GGAAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700



Db 2742 GAGGCCGCTGAGGCCACCGGCCGAGAGGGCGCTCAAAGCAAGAAAGAGAGGGCGGGAG 2801  
Qy 2701 CGGCAAGGGCCACCACTGCCAAGAGCTCGGGCGCCCCCGAGGACAGGACTCCAGTGTAC 2760  
Db 2802 CGGCAAGGGCCACCAAGGCCAGAGGCTCGGGCGCCCCCGAGGACAGGACTCCAGTGTCCAC 2861  
Qy 2761 CTGCAAGTGCAGACGAGGTGGATGAGGCCGGAAGGCGGCGGCAAGAAACCGGCTGTGCCCC 2820  
Db 2862 CTGCAAGTGCAGACGAGGTGGATGAGGCCGGAAGGCGGCGGCAAGAAACCGGCTGTGCCCC 2921  
Qy 2821 AAGGCCGAGCTCTCTCAACCCGCACTGGAGACCCCGGGGCAATGACTCTACCCCAAGAGCC 2880  
Db 2922 AAGGCCGAGCTCTCTCAACCCGCACTGGAGACCCCGGGGCAATGACTCTACCCCAAGAGCC 2981  
Qy 2881 ACTGGAAGCTGAAGAGCTGAAGCAAGCAAGCGAGCTGCATCCCCCATCCAGTCAACAA 2940  
Db 2982 ACTGGAAGCTGAAGAGCTGAAGCAAGCGAGCGGCTGCATCCCCCATCCAGTCAACAA 3041  
Qy 2941 AGTCCATGAGCCCCCGGGAAGAGCGAGCTCCACCAAGCGCTCCCCAGCCCCCACC 3000  
Db 3042 AGTCCATGAGCCCCCGGGAAGAGCGAGCTCCACCAAGCGCTCCCCAGCCCCCACC 3101  
Qy 3001 GCCACCGCAAAACCTGCAAGCGGAGAGCGAGCGCTTCAGCAAGCTTCGAGAGCGCCCCG 3060  
Db 3102 GCCACCGCAAAACCTGCAAGCGGAGAGCGAGCGCTTCAGCAAGCTTCGAGAGCGCCCCG 3161  
Qy 3061 GGGCAAGAGCAGAGAGCCCGGCAACCCCGCGCGACAAAGAG----- 3100  
Db 3162 GGGCAAGAGCAGAGAGCCCGGCAACCCCGCGCGACAAAGAGAGAGCTGTGTTT 3221  
Qy 3101 ----GCTTTCGAGGCCGAGGCCCAAGAGCTGCTGGGAGCCCCCTTGTGAGATTTCGG 3156  
Db 3222 CCGACCTTCGAGCGCGAGGCCAAGAGCTGCTGGGAGCCCCCTTGTGAGATTTCGG 3281  
Qy 3157 CCGACCTTCGAGCGCGAGGCCAAGAGCTGCTGGGAGCCCCCTTGTGAGATTTCGG 3216  
Db 3282 CCGACCTTCGAGCGCGAGGCCAAGAGCTGCTGGGAGCCCCCTTGTGAGATTTCGG 3341  
Qy 3217 CTCAAGCTTCTCTTAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAG 3276  
Db 3342 CTCAAGCTTCTCTTAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAG 3401  
Qy 3277 CCGGCGGCTCTGCGCGCGCGCCCAACCTTCACACCGGCTCTCTCTCTCTCTCTCTCT 3336  
Db 3402 CCGGCGGCTCTGCGCGCGCGCCCAACCTTCACACCGGCTCTCTCTCTCTCTCTCTCT 3461  
Qy 3337 CAAGCAACCCAGCGCTCTGAGAGGCAATAGAGTGCATCTCCAAAGAAATGCTGTCCA 3396  
Db 3462 CAAGCAACCCAGCGCTCTGAGAGGCAATAGAGTGCATCTCCAAAGAAATGCTGTCCA 3521  
Qy 3397 GCTCCAGCTCCGTAAGCTCAAGATGCCAAGGCCCGGCTGAGCGCTGTCAACATGGGCT 3456  
Db 3522 GCTCCAGCTCCGTAAGCTCAAGATGCCAAGGCCCGGCTGAGCGCTGTCAACATGGGCT 3581  
Qy 3457 GCGGCTGCTCATGAGACCCCAAAAGCTGGCACTTCAGCGGAGTGAACAGAGCAAGCT 3516  
Db 3582 GCGGCTGCTCATGAGACCCCAAAAGCTGGCACTTCAGCGGAGTGAACAGAGCAAGCT 3641  
Qy 3517 GTCGCCAGGGGCGAGGCTGGGCGACCGGAGAGCTGAGGGGTGGCCACAGGCCAAGAGGC 3576  
Db 3642 GTCGCCAGGGGCGAGGCTGGGCGACCGGAGAGCTGAGGGGTGGCCACAGGCCAAGAGGC 3701  
Qy 3577 GTCGCTGCTGAGAGGAGAGCTCTGGGCTCAAGTTCGGGCGGAGCAATCAACAAAGCAT 3636  
Db 3702 GTCGCTGCTGAGAGGAGAGCTCTGGGCTCAAGTTCGGGCGGAGCAATCAACAAAGCAT 3761  
Qy 3637 TCCCAAGCAAGGGGTGCGCTTGGAGCAAGCGCATCAATACGGGCTTCATCAACCCACG 3696  
Db 3762 TCCCAAGCAAGGGGTGCGCTTGGAGCAAGCGCATCAATACGGGCTTCATCAACCCACG 3821  
Qy 3697 CACGCAAGCTAAGCTCTGTACAAAGGAGCAATCAACAGAGTATCGGCGAGAGAGAGCC 3756  
Db 3822 CACGCAAGCTAAGCTCTGTACAAAGGAGCAATCAACAGAGTATCGGCGAGAGAGAGCC 3881

Qy 3757 GAGTGCCTTGAACCGCGGCGGGAGAGAGCTGTGCCAAAGGCGCAAGTATCTACGAAG 3816  
Db 3882 GAGTGCCTTGAACCGCGGCGGGAGAGAGAGCTGTGCCAAAGGCGCAAGTATCTACGAAG 3941  
Qy 3817 CAAGAAAGGCCACGCTTGTCTTATGAGGTTGAGATGTCTGTGACCCAGTCTCCAGGA 3876  
Db 3942 CAAGAAAGGCCACGCTTGTCTTATGAGGTTGAGATGTCTGTGACCCAGTCTCCAGGA 4001  
Qy 3877 GAGCGGCAAGAGCAGCTCAGGACCCCGCCATGAGACGGCGGCGCCCAAGCGCACTATGA 3936  
Db 4002 GAGCGGCAAGAGCAGCTCAGGACCCCGCCATGAGACGGCGGCGCCCAAGCGCACTATGA 4061  
Qy 3937 CATGATGAGAGGCGCGGTGGGAGAGCCATCTCTCAAGCAGCATGAAAGTCTCATGGG 3996  
Db 4062 CATGATGAGAGGCGCGGTGGGAGAGCCATCTCTCAAGCAGCATGAAAGTCTCATGGG 4121  
Qy 3997 CCGTGCATTCGCGCGGAGAGCGACAGGCCCCCACTCAAAAGAGACACATCCG 4056  
Db 4122 CCGTGCATTCGCGCGGAGAGCGACAGGCCCCCACTCAAAAGAGACACATCCG 4181  
Qy 4057 CCGGTCAATCAACAAGGGATTCCTCGGTCTTACGTGAGGCAAGAGGACTTACTGTG 4116  
Db 4182 CCGGTCAATCAACAAGGGATTCCTCGGTCTTACGTGAGGCAAGAGGACTTACTGTG 4241  
Qy 4117 TCGGAGGCGCAAGCTCTTAAAGCGGAGGCGACGCTTCGCGCCCAACCGCTCAAGG 4176  
Db 4242 TCGGAGGCGCAAGCTCTTAAAGCGGAGGCGACGCTTCGCGCCCAACCGCTCAAGG 4301  
Qy 4177 CCTGACCGAGGCTTAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4236  
Db 4302 CCTGACCGAGGCTTAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4361  
Qy 4237 GGGAGCTGTGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4296  
Db 4362 GGGAGCTGTGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4421  
Qy 4297 GCTGAGGAGCAAGCGCGGAGCTGCGCTTGGCCCGGCGGCGCTCAAGAGGAGCTCATAC 4356  
Db 4422 GCTGAGGAGCAAGCGCGGAGCTGCGCTTGGCCCGGCGGCGCTCAAGAGGAGCTCATAC 4481  
Qy 4357 GCAAGGAGCAACCGGCTCAAGTACAGACCGGCGGCTTCACACCTGGCTCAAAAGAGCA 4416  
Db 4482 GCAAGGAGCAACCGGCTCAAGTACAGACCGGCGGCTTCACACCTGGCTCAAAAGAGCA 4541  
Qy 4417 GGTACGCTCCCTCATGAGAGCCCGGCGGAGCGTTCACACCGGAGCAACCGGCTGATGT 4476  
Db 4542 GGTACGCTCCCTCATGAGAGCCCGGCGGAGCGTTCACACCGGAGCAACCGGCTGATGT 4601  
Qy 4477 GATGAGCGAGCGCGGCGGAGCTGGAACGTGCTGTACAGAGAGAGCTGGAAGAGCGGCG 4536  
Db 4602 GATGAGCGAGCGCGGCGGAGCTGGAACGTGCTGTACAGAGAGAGCTGGAAGAGCGGCG 4661  
Qy 4537 AGGAGCGGCGAGAGCTGAGGAGGAGCTCAATGCGGCGGCGCGGAGCTTGTGCTGGA 4596  
Db 4662 AGGAGCGGCGAGAGCTGAGGAGGAGCTCAATGCGGCGGCGCGGAGCTTGTGCTGGA 4721  
Qy 4597 GCTGGGTAAAGCGGAGGAGCGGCGGAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 4656  
Db 4722 GCTGGGTAAAGCGGAGGAGCGGCGGAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 4781  
Qy 4657 CCACCTCCAGAGAGTTCGCGGAGCAATGCGGAGGCGGAGGCGGAGGCGGAGGAGGAGG 4716  
Db 4782 CCACCTCCAGAGAGTTCGCGGAGCAATGCGGAGGCGGAGGCGGAGGCGGAGGAGGAGG 4841  
Qy 4717 CAGGCTTTCGTTCAGCAAGAGCAATCCAGAGCCGAAGCTGACGCTCGTGAAT 4776  
Db 4842 CAGGCTTTCGTTCAGCAAGAGCAATCCAGAGCCGAAGCTGACGCTCGTGAAT 4901  
Qy 4777 CGGCAAGTCCCGGAGAGCAAGCAAGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4836  
Db 4902 CGGCAAGTCCCGGAGAGCAAGCAAGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4961

QY 4837 GCACTGTTCGGGGCGGTGAGTGGCTGGA CTTGTATCGAGCCACATCCCTGTGCTT 4896  
DB GCACTGTTCGGGGCGGTGAGTGGCTGGA CTTGTATCGAGCCACATCCCTGTGCTT 5021  
QY 4897 GCACCCCCATCTCAATACCCCGGGCATCTCTGTGAGCGAGCCGCTGCTTACTACCTGCC 4956  
DB 5022 GCACCCCCATCTCAATACCCCGGGCATCTCTGTGAGCGAGCCGCTGCTTACTACCTGCC 5081  
QY 4957 CGGACACTGGAGCCCGCAACCCCACTTACCCGCA CTTGTACCCCACTTCACTCATCCGCG 5016  
DB 5082 CGGACACTGGAGCCCGCAACCCCACTTACCCGCA CTTGTACCCCACTTCACTCATCCGCG 5141  
QY 5017 CTACCCCGACACGGCGGGCTGTGAGAACCGGAGACCATCATCATATGACTTACATCTTC 5076  
DB 5142 CTACCCCGACACGGCGGGCTGTGAGAACCGGAGACCATCATCATATGACTTACATCTTC 5201  
QY 5077 GCGAGCATGTGACACACACACCGGCAACCGGCACTGAGCCGAGAGCTGTATTCCTGAGGG 5136  
DB 5202 GCGAGCATGTGACACACACACCGGCAACCGGCACTGAGCCGAGAGCTGTATTCCTGAGGG 5261  
QY 5137 CCTCTGCCCCGGGAGTCTCGGTGACCTCAACTACGCTGCGGGTCCCGAGGACATCAT 5196  
DB 5262 CCTCTGCCCCGGGAGTCTCGGTGACCTCAACTACGCTGCGGGTCCCGAGGACATCAT 5321  
QY 5197 CGACTGTCTCCAAATGACACACTGCTGTGTCTGTGCCCCCGACACAGGCAACCCGAGC 5256  
DB 5322 CGACTGTCTCCAAATGACACACTGCTGTGTCTGTGCCCCCGACACAGGCAACCCGAGC 5381  
QY 5257 CACCGCATGTGACCGGCTTGTCTTACTTCCCAACCGGCGCCGAGCCTTTCAGAGCCGCA 5316  
DB 5382 CACCGCATGTGACCGGCTTGTCTTACTTCCCAACCGGCGCCGAGCCTTTCAGAGCCGCA 5441  
QY 5317 CAGCACTCCCACTCTCCCCAGAGGTCCAGACACTTGTGACAAACCAACCAACCAAGTC 5376  
DB 5442 CAGCACTCCCACTCTCCCCAGAGGTCCAGACACTTGTGACAAACCAACCAACCAAGTC 5501  
QY 5377 CTCTGTCCGAGCGGAGCGAGACCGGAGTCGAGAGCGGAGACCGGAGTCGAGAGCGGAGAA 5436  
DB 5502 CTCTGTCCGAGCGGAGCGAGACCGGAGTCGAGAGCGGAGACCGGAGTCGAGAGCGGAGAA 5561  
QY 5437 GTTCATCTCTCAAGTCCACACAGAGGTGAGACACGACCTTGTGAGACTTGTGTACGA 5496  
DB 5562 GTTCATCTCTCAAGTCCACACAGAGGTGAGACACGACCTTGTGAGACTTGTGTACGA 5621  
QY 5497 GCGAGCAAGGTGAGAGCGGAGCGGAGCGGCGGGGGGGGGGGGAGAGCAAGCGGCCGCG 5556  
DB 5622 GCGAGCAAGGTGAGAGCGGAGCGGAGCGGCGGGGGGGGGGGGAGAGCAAGCGGCCGCG 5681  
QY 5557 CTCTCACTCCCATGCCCCACAGCACTCGCCCATCTCCCTCGGACCCAGAGTCCCTTCA 5616  
DB 5682 CTCTCACTCCCATGCCCCACAGCACTCGCCCATCTCCCTCGGACCCAGAGTCCCTTCA 5741  
QY 5617 GCGAGACCCCAAGTGTGCTTCAACACAGGATGAAAGGTATCATCCGCTGTGAGGCC 5676  
DB 5742 GCGAGACCCCAAGTGTGCTTCAACACAGGATGAAAGGTATCATCCGCTGTGAGGCC 5801  
QY 5677 CAGCAAGCCCAAGGTCTGAGGTCTCAACTTCAACTCTTCAACCGGTCGCGCACTGTGCAC 5736  
DB 5802 CAGCAAGCCCAAGGTCTGAGGTCTCAACTTCAACTCTTCAACCGGTCGCGCACTGTGCAC 5861  
QY 5737 ATTCCCACTGACACCACTGACCACTGGGCGGACCTTGTGAGGGTCTTACCTTACCT 5796  
DB 5862 ATTCCCACTGACACCACTGACCACTGGGCGGACCTTGTGAGGGTCTTACCTTACCT 5921  
QY 5797 CATGAGCCCGTCTTGTCTGCCCCAAGAGAGCGCCCGGGTGCCTCGGACAGAGCGGCCG 5856  
DB 5922 CATGAGCCCGTCTTGTCTGCCCCAAGAGAGCGCCCGGGTGCCTCGGACAGAGCGGCCG 5981  
QY 5857 AGAGAGACCGGCGCAAGCTTCTTCCGCAAGCGGCCCAAGCCCGCTCGGAGGTGAGAGCCG 5916  
DB 5982 AGAGAGACCGGCGCAAGCTTCTTCCGCAAGCGGCCCAAGCCCGCTCGGAGGTGAGAGCCG 6041  
QY 5917 CTCTCTCCCAAGCAAGGCTCGAGAGCCCGGCGCTTGTGTCTTGTGCGCACGCG 5976

DB 6042 CTCTCTCCCAAGCAAGGCTCGAGAGCCCGGCGCTTGTGTCTTGTGCGCACGCG 6101  
QY 5977 CACCATGAGCCCGGACCCCTGCGAAGAACCTGACCTTCAACACAGCCGAGCCCGGACCGCC 6036  
DB 6102 CACCATGAGCCCGGACCCCTGCGAAGAACCTGACCTTCAACACAGCCGAGCCCGGACCGCC 6161  
QY 6037 GCGGCAACTGCTCGGCTCGAGCCCGGACCCGGAAGAACTCAAGTAAACCTTTTC 6096  
DB 6162 GCGGCAACTGCTCGGCTCGAGCCCGGACCCGGAAGAACTCAAGTAAACCTTTTC 6221  
QY 6097 CATTCAGAACTGAACTTCCGTTCTTGTGGTTTACAGCGGAGAGCTTACAGCCCGAAG 6156  
DB 6222 CATTCAGAACTGAACTTCCGTTCTTGTGGTTTACAGCGGAGAGCTTACAGCCCGAAG 6281  
QY 6157 GGTGAGCCCGTCAAGCCCTGTGAGCTCACCACTGTGACCCAGACAGAGGGGCTCCCA 6216  
DB 6282 GGTGAGCCCGTCAAGCCCTGTGAGCTCACCACTGTGACCCAGACAGAGGGGCTCCCA 6341  
QY 6217 GCACTGTGAAAGCTCGACAGAGCACTGTGAGGGGAGCTGCGGCCAAGAGCGCAG 6276  
DB 6342 GCACTGTGAAAGCTCGACAGAGCACTGTGAGGGGAGCTGCGGCCAAGAGCGCAG 6401  
QY 6277 CCCCCTGAAAGCTTGTGCGGGAGGCGCCCACTTCCACACTGCGGCGCTGCTTGAAG 6336  
DB 6402 CCCCCTGAAAGCTTGTGCGGGAGGCGCCCACTTCCACACTGCGGCGCTGCTGAAG 6461  
QY 6337 CGAGCCCTGTCCAGGCGCGCTGCTCGAGACCGGCCAGGGGTCAAGTACAGGGGGT 6396  
DB 6462 CGAGCCCTGTCCAGGCGCGCTGCTCGAGACCGGCCAGGGGTCAAGTACAGGGGGT 6521  
QY 6397 GGTCACTCTGAGCCCAAGACATCATGAGGTATCATACAGAACTTACACCCGAGCACCC 6456  
DB 6522 GGTCACTCTGAGCCCAAGACATCATGAGGTATCATACAGAACTTACACCCGAGCACCC 6581  
QY 6457 ACAGAGCTTCAAGGCAACCCCTGCGGCCCCCTTACTCTCTTCTGTGGGCGAGCTGCC 6516  
DB 6582 ACAGAGCTTCAAGGCAACCCCTGCGGCCCCCTTACTCTCTTCTGTGGGCGAGCTGCC 6641  
QY 6517 CGTCTGTGACCTCCGCGGCGCACCCAGTACCTTTCCTCCGCGCCCGGACCAATGGTGC 6576  
DB 6642 CGTCTGTGACCTCCGCGGCGCACCCAGTACCTTTCCTCCGCGCCCGGACCAATGGTGC 6701  
QY 6577 CCGGCGCCGTGCTCCCGCACAGCGAAGGGGCAAGGTTCTCAGAGCCAAACAGAC 6636  
DB 6702 CCGGCGCCGTGCTCCCGCACAGCGAAGGGGCAAGGTTCTCAGAGCCAAACAGAC 6761  
QY 6637 GTCTGTCTTGTGGTGTGTGAGGACGGTATTGAACCTGTGTCCACCGAGGGCATGAC 6696  
DB 6762 GTCTGTCTTGTGGTGTGTGAGGACGGTATTGAACCTGTGTCCACCGAGGGCATGAC 6821  
QY 6697 GAGAGCCAGGCACTCCCGGAGTGTGTATCCCGCTGTGTATCCGGAGTGGGGAACAAGAC 6756  
DB 6822 GAGAGCCAGGCACTCCCGGAGTGTGTATCCCGCTGTGTATCCGGAGTGGGGAACAAGAC 6881  
QY 6757 GAGAGCCAGGAGATGGGCTCCAAAGTCTCAGGCAACACAGGCAAGCCGACCTTCTT 6816  
DB 6882 GAGAGCCAGGAGATGGGCTCCAAAGTCTCAGGCAACACAGGCAAGCCGACCTTCTT 6941  
QY 6817 CAGCAAGCTGACCGAGAGCACTCGGCATGTGCAAGTCCAGAGCAAGAGTCAACAA 6876  
DB 6942 CAGCAAGCTGACCGAGAGCACTCGGCATGTGCAAGTCCAGAGCAAGAGTCAACAA 7001  
QY 6877 GAAAGCTGAACCCCAACACCGAATGAGCTGAATTCAATTATGACGACTGGGAGCGA 6936  
DB 7002 GAAAGCTGAACCCCAACACCGAATGAGCTGAATTCAATTATGAGCAAGCTGGGAGCGA 7061  
QY 6937 GATCTTCAATTATGCGCGCATCACCGGAAACAGGCTTATGACTTATGAAAGCCAGCGGT 6996  
DB 7062 GATCTTCAATTATGCGCGCATCACCGGAAACAGGCTTATGACTTATGAAAGCCAGCGGT 7121  
QY 6997 GCGAGAACTGCCAGACCAACATGGGAGCTGAGGCGCATATTGAAAGGCACTCATGG 7056

Db	7122	CCAGGAACTGCGCAGACCAACATG9GGGCTG9AGGGCCATTAATTGAAAGGACCTCATG9G	7181
QY	7057	TAAATATATACAGATGGGAAAGTGTCCCGGCGCTGACGGCCAAATGCTTTTAAACCTCTGAA	7116
Db	7182	TAAATATATACCAAGTGGGAAGAGTCCCGCCCGCTGAGCCGCATATGCTTTTAAACCTCTGAA	7241
QY	7117	TGCGAGTGTCCAGGCGCGCGCGGTGCTATGCGCATPAACGCTGTGACGGAAGGAAATGACCA	7176
Db	7242	TGCGAGTGTCCAGGCGCGCGCGGTGCTATGCGCATPAACGCTGTGACGGAAGGAAATGACCA	7301
QY	7177	CACACTCACCTCGCCAGGTGGCGGGGAAAGGACAAAGTCTCTGGCAGAACCCAGACGCG	7236
Db	7302	CACACTCACCTCGCCAGGTGGCGGGGAAAGGACAAAGTCTCTGGCAGAACCCAGACGCG	7361
QY	7237	AAAAAGCCAGTGTCCCGGCGCGCGCTTGAGCATCTGGGAGACGAGCCACCTCTGTCTCTCTC	7296
Db	7362	AAAAAGCCAGTGTCCCGGCGCGCGCTTGAGCATCTGGGAGACGAGCCACCTCTGTCTCTCTC	7421
QY	7297	AGTGCACCTCGGAGGGAGATGTGCACACCGCGGAGACGCGCTGCACCAACCGCGTGTGGAGGA	7356
Db	7422	AGTGCACCTCGGAGGGAGATGTGCACACCGCGGAGACGCGCTGCACCAACCGCGTGTGGAGGA	7481
QY	7357	CAGGCGCCTCGTCCGCGAGGTTCACAGCCATTCCCTCAACAACCCCTGATCATGCGCTGTCA	7416
Db	7482	CAGGCGCCTCGTCCGCGAGGTTCACAGCCATTCCCTCAACAACCCCTGATCATGCGCTGTCA	7541
QY	7417	GGCGGGTGTCAATGCGCTTCCCAACCCCGACCGGCGCTCCCGCGGAGCAGCGGCGCCCTTGC	7476
Db	7542	GGCGGGTGTCAATGCGCTTCCCAACCCCGACCGGCGCTCCCGCGGAGCAGCGGCGCCCTTGC	7601
QY	7477	TGGCGCCCAACACCGCTGGGAGACGAGAGGCCAAGCATGTGCTGTGTGCGAGTACAGAC	7536
Db	7602	TGGCGCCCAACACCGCTGGGAGACGAGAGGCCAAGCATGTGCTGTGTGCGAGTACAGAC	7661
QY	7537	ACTCTCCGACAGCGAGTGACTCAGAACAGGCGGGGGGGGGCGGGCGGTBTUBERPERLUG	7596
Db	7662	ACTCTCCGACAGCGAGTGACTCAGAACAGGCGGGGGGGGGCGGGCGGTBTUBERPERLUG	7711
QY	7597	TCAGGTCCAGCGAGCCACAGAAACGAGCCCTGACAGAGCGGGGCGCGGTGCGCATCCGCC	7656
Db	7712	TCAGGTCCAGCGAGCCACAGAAACGAGCCCTGACAGAGCGGGGCGCGGTGCGCATCCGCC	7711
QY	7657	AACCAAGGAAGGAACCCCTTAGTCCGCTGGGCTTCCATCTGTCTGTCTCCAGAGCG	7716
Db	7772	AACCAAGGAAGGAACCCCTTAGTCCGCTGGGCTTCCATCTGTCTGTCTCCAGAGCG	7831
QY	7717	GCATCTCTTGGCTGTCTTAAAGCCTTAACATPAAGCTCCCGGCGTGGGCGCTGTGACGA	7776
Db	7832	GCATCTCTTGGCTGTCTTAAAGCCTTAACATPAAGCTCCCGGCGTGGGCGCTGTGACGA	7891
QY	7777	CCTTACTCAGGGAGATGTTTACCTGTGTGCTGAGGAAGGAGAGGGAAGGAGCGGAGAGGAG	7836
Db	7892	CCTTACTCAGGGAGATGTTTACCTGTGTGCTGAGGAAGGAGAGGGAAGGAGCGGAGAGGAG	7951
QY	7837	GCAAGGCGAGGCGTGTGGAGCCACACAGGCGGCGAGGCGGCGCAGGGAACCCAAAGCAG	7896
Db	7952	GCAAGGCGAGGCGTGTGGAGCCACACAGGCGGCGAGGCGGCGCAGGGAACCCAAAGCAG	8011
QY	7897	GATGACCAACGACCTCCACAGCCACATGCGCTCCCGGAAATGACATTTTGGAAACAAAGCTTAA	7956
Db	8012	GATGACCAACGACCTCCACAGCCACATGCGCTCCCGGAAATGACATTTTGGAAACAAAGCTTAA	8071
QY	7957	CTGAGCTTCAGGCGCGCGCGCTTCCCTCCGCTCCCATCCCGCTTAAAGCGCTGTGACAG	8016
Db	8072	CTGAGCTTCAGGCGCGCGCGCTTCCCTCCGCTCCCATCCCGCTTAAAGCGCTGTGACAG	8131
QY	8017	ATGAGACGAGGCGCTGTTCAGGCGCCCAATGTGCGCTGTTCGCGTCCCAACAGCTGCCCA	8076
Db	8132	ATGAGACGAGGCGCTGTTCAGGCGCCCAATGTGCGCTGTTCGCGTCCCAACAGCTGCCCA	8191
QY	8077	GCCAAACGAGATTGCTGGAACCAACATCAGGCGCAGGTGGCGGACAAAGGGCGCAGTGTGC	8136
Db	8192	GCCAAACGAGATTGCTGGAACCAACATCAGGCGCAGGTGGCGGACAAAGGGCGCAGTGTGC	8251

OY	8137	GCCTGGGGGGAA	CGGATGCTCCAGAGACTGGA	CTGTTTCTTTTTCACATCTGTGCGGAG	8196
Db	8252	GCCGTGGGGGAA	CGGATGCTCCAGAGACTGGA	CTGTTTCTTTTTCACATCTGTGCGGAG	8311
OY	8197	CGGTGGGAAAGAA	AGCGAGATGTAATGATGTGTGGTTTACAGGGTAATATTTTGATAC		8256
Db	8312	CGGTGGGAAAGAA	AGCGAGATGTAATGATGTGTGGTTTACAGGGTAATATTTTGATAC		8371
OY	8257	CTTCATGAATTAAT	TTCAGATGTTTTCACGAGAGAGACTTACCCAGTATTA	CTGCTGC	8316
Db	8372	CTTCATGAATTAAT	TTCAGATGTTTTCACGAGAGAGACTTACCCAGTATTA	CTGCTGC	8431
OY	8317	TGTGCTTTTGAT	CTCTGCTTACCGTTCAAGAGCGTGTGACAGGCCGACAGTCCGTGACCC		8376
Db	8432	TGTGCTTTTGAT	CTCTGCTTACCGTTCAAGAGCGTGTGACAGGCCGACAGTCCGTGACCC		8491
OY	8377	CATACACTGACGAG	CAAGGGGGGGGAGCTGCTGTCACGCGCCGCTGATCCTCCCTC		8436
Db	8492	CATACACTGACGAG	CAAGGGGGGGGAGCTGCTGTCACGCGCCGCTGATCCTCCCTC		8551
OY	8437	CCCTCCCTTCTCT	TGGGCGAATGAATTCGATTCGATTCGTGAGCCGCGCATTTGGCGAGG		8496
Db	8552	CCCTCCCTTCTCT	TGGGCGAATGAATTCGATTCGATTCGTGAGCCGCGCATTTGGCGAGG		8611
OY	8497	TGTGTGTAATTC	TGTCATTTTACACACGTCGTTCTAATTA	AAAAAGCAATTAATCTCCAAA	8556
Db	8612	TGTGTGTAATTC	TGTCATTTTACACACGTCGTTCTAATTA	AAAAAGCAATTAATCTCCAAA	8671
OY	8557	AAAAAAAAAAAAAA	8571		
Db	8672	AAAAAAAAAAAAAA	8686		
RESULT 9					
ADG86290					
ID	ADG86290	standard; cDNA; 8686 BP.			
XX	AC	ADG86290;			
XX	DT	11-MAR-2004 (first entry)			
DE	Human SMRT	encoding cDNA SEQ ID NO:4.			
XX	KW	SMRT; silencing mediator for retinoid and thyroid hormone action;			
KW	SMRT inhibitor; cytosolic; antiinflammatory; antiarthritic;				
KW	antirheumatic; antisense therapy; inflammatory disorder;				
KW	rheumatoid arthritis; hyperproliferative disorder; cancer; leukaemia;				
XX	breast cancer; human; gene; ss.				
XX	OS	Homo sapiens.			
XX	XX				
FH	Key	Location/Qualifiers			
FT	CDS	157..7680			
FT	FT	/*cag= a			
FT	FT	/product= "SMRT"			
XX	PN	MO2003106645-A2.			
XX	PD	24-DEC-2003.			
XX	XX				
PF	17-JUN-2003;	2003WO-US018923.			
XX	XX				
PR	17-JUN-2002;	2002US-00174014.			
XX	XX				
PA	(ISIS-)	ISIS PHARM INC.			
XX	XX				
PI	Bennett CF, Freier SM, Dobie KM;				
XX	XX				
DR	WPI; 2004-082184/08.				
DR	P-PSDB; ADG86291.				
DR	GENBANK; AF125672.				
XX	XX				

PT Novel antisense compound targeted to nucleic acid encoding SMRT  
PT (silencing mediator for retinoid and thyroid hormone action), useful for  
PT treating animal having disease associated with SMRT such as cancer,  
PT rheumatoid arthritis.

XX Example 13; SEQ ID NO 4; 260pp; English.

CC The present invention describes a compound (I) 8-50 nucleobases in length  
CC targeted to a nucleic acid molecule encoding SMRT (silencing mediator for  
CC retinoid and thyroid hormone action), where (I) specifically hybridises  
CC with the nucleic acid molecule encoding SMRT and inhibits expression of  
CC SMRT. (I) specifically hybridises with at least 8-nucleobase portion of a  
CC preferred target region on nucleic acid molecule encoding SMRT. Also  
CC described is a composition (II) comprising (I) and a carrier or diluent.  
CC (I) and (II) have cytostatic, antiinflammatory, antiarthritic and  
CC antirheumatic activities, and can be used in antisense therapy, and as  
CC SMRT expression inhibitors. (I) is useful for inhibiting the expression  
CC of SMRT in cells or tissues. (I) is also useful for treating an animal  
CC having a disease or condition associated with SMRT, e.g., inflammatory  
CC disorder such as rheumatoid arthritis; or a hyperproliferative disorder  
CC such as cancer chosen from leukemia and breast cancer, by inhibiting the  
CC expression of SMRT. (I) is useful for diagnostics, therapeutics,  
CC prophylaxis and as research reagents and kits. The present sequence  
CC encodes human SMRT, which is used in an example from the present  
CC invention.

SO Sequence 8686 BP; 1881 A; 3069 C; 2562 G; 1174 T; 0 U; 0 Other;

Query Match 97.2%; Score 8324.6; DB 12; Length 8686;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 8472; Conservative 1; Mismatches 34; Indels 88; Gaps 4;

QY 1 CATTGCTGGGCTCCACACAGCTTGTGGACAGAGTGGAGGCGCACTGAGCCCGCTAACCC 60  
DB 156 CATTGCTGGGCTCCACACAGCTTGTGGACAGAGTGGAGGCGCACTGAGCCCGCTAACCC 215  
QY 61 GCCCCACAGCCTTTCTTACCAGTGCAGATCGCCCGACGACAGCGACGTCGGGCTCCT 120  
DB 216 GCCCCACAGCCTTTCTTACCAGTGCAGATCGCCCGACGACAGCGACGTCGGGCTCCT 215  
QY 121 GGAGTACACAGACCACTCCCGGAGTATGCTTCCACCTGTGCGCGGCTCATCATCCA 180  
DB 276 GGAGTACACAGACCACTCCCGGAGTATGCTTCCACCTGTGCGCGGCTCATCATCCA 335  
QY 181 GCCCAGCGCGGAGGCGCTCCCTGCTGTGATGTTCCAGCCCGGAAATGAAGGTTCCA 240  
DB 336 GCCCAGCGCGGAGGCGCTCCCTGCTGTGATGTTCCAGCCCGGAAATGAAGGTTCCA 395  
QY 241 GGAGCTCCACTGCGGCGACAGTCCCACTCATACCTGCGGAGCTGGGAAATGCAGAGAT 300  
DB 396 GGAGCTCCACTGCGGCGACAGTCCCACTCATACCTGCGGAGCTGGGAAATGCAGAGAT 455  
QY 301 GGAGTTCAATTGAAGAGACGCGCTCGGCTAGAGCTGCTGACCTCCCTGCTGACCC 360  
DB 456 GGAGTTCAATTGAAGAGACGCGCTCGGCTAGAGCTGCTGACCTCCCTGCTGACCC 515  
QY 361 GTCACTCCCTGCTGCGGCGACGCGCTGCGGAGATTTGAAGACTCAACCAAGACCGTGA 420  
DB 516 GTCACTCCCTGCTGCGGCGACGCGCTGCGGAGATTTGAAGACTCAACCAAGACCGTGA 575  
QY 421 CTTGACGGGCAAGCTGGAACCGGTGTCTCCCCCAGCGCCCGGACACATGACCTTAGACT 480  
DB 576 CTTGACGGGCAAGCTGGAACCGGTGTCTCCCCCAGCGCCCGGACACATGACCTTAGACT 635  
QY 481 GGAGCTGTGCGGCGACGCGCTGTCCAAGAGAGAGCTGATCCAGAACATGACCGCGTGA 540  
DB 636 GGAGCTGTGCGGCGACGCGCTGTCCAAGAGAGAGCTGATCCAGAACATGACCGCGTGA 695  
QY 541 CCGAGAGATCAACATGTGTAGAGAGAGATCTTAAAGCTGAAGAAAGAGAGAGAGAGCT 600  
DB 696 CCGAGAGATCAACATGTGTAGAGAGAGATCTTAAAGCTGAAGAAAGAGAGAGAGAGCT 755  
QY 601 GGAGAGAGAGCTGCAAGCGCGCGAGCTGAGAAAGCCGTTGTCAACCGCGCCCATCGA 660

DB 756 GGAGAGAGAGCTGCCAAGCGCCCGACCTGGAAGCCCGTGTCAACCGCGCCCATCGA 815  
QY 661 GTGGAAGACCCGAGCGCTGTGCAGATCATCTACGACGAGAACCGGAAGAGGCTGAAGC 720  
DB 816 GTGGAAGACCCGAGCGCTGTGCAGATCATCTACGAGAACCGGAAGAGGCTGAAGC 875  
QY 721 TGCACATCCGATTTCTGAAAGGCTTGAGGCGCCGAGTGAAGCTCCGCTGTACACACGCC 780  
DB 876 TGCACATCCGATTTCTGAAAGGCTTGAGGCGCCGAGTGAAGCTCCGCTGTACACACGCC 935  
QY 781 CTCGACACCCGCGAGTATCATGAGAACATCAAAATPAACGAGCGATGCGAAGAACT 840  
DB 936 CTCGACACCCGCGAGTATCATGAGAACATCAAAATPAACGAGCGATGCGAAGAACT 995  
QY 841 AATCTGTACTTCAAGAGAGAGAAATCAAGCTCGAAGAACATGAGACAGAAATTTGCGCA 900  
DB 996 AATCTGTACTTCAAGAGAGAGAAATCAAGCTCGAAGAACATGAGACAGAAATTTGCGCA 1055  
QY 901 GCGCTATGACCAAGCTCATGAGAGGCTTTGAAAAAAGTGAAGCGCATCGAAAAAACACC 960  
DB 1056 GCGCTATGACCAAGCTCATGAGAGGCTTGGAAGAAAGTGAAGCGCATCGAAAAAACACC 1115  
QY 961 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGAGTACTACGAAAAAGCACTTCCCTGAGAT 1020  
DB 1116 CCGGCGCGGCGCAAGAGAGCAAGGTTCGCGAGTACTACGAAAAAGCACTTCCCTGAGAT 1175  
QY 1021 CCGGACAGACCGGAGACTCATGAGAGCGCATGACAGACAGGTGTGGCCAGCGGCGAGTGG 1080  
DB 1176 CCGGACAGACCGGAGACTCATGAGAGCGCATGAGAGCGCATGAGAGCGCATGAGAGAGTGG 1232  
QY 1081 GCTGTCCATGTGCGGCGCGCCGAGAGACGAGAGTGTGAGATCATGATGAGGCTCTC 1140  
DB 1233 GCTGTCCATGTGCGGCGCGCCGAGAGACGAGAGTGTGAGATCATGATGAGGCTCTC 1292  
QY 1141 AGAGCAGAGAAACTTGAAGAGAGAGATGCGCAGCTGCGCTGATCCCGCCATGCTGTGA 1200  
DB 1293 AGAGCAGAGAAACTTGAAGAGAGAGATGCGCAGCTGCGCTGATCCCGCCATGCTGTGA 1352  
QY 1201 CGAGCTGACACAGACCGCATCAAGTTTCATCAACATGAAACGCGCTTATGCGGACCCCAT 1260  
DB 1353 CGAGCTGACACAGACCGCATCAAGTTTCATCAACATGAAACGCGCTTATGCGGACCCCAT 1412  
QY 1261 GAAAGTGTCAAAAGACCGCAGGTCATGAACATGTGAAGTGAAGAGAGAGAGACTT 1320  
DB 1413 GAAAGTGTCAAAAGACCGCAGGTCATGAACATGTGAAGTGAAGAGAGAGAGACTT 1472  
QY 1321 CCGGAGAAAGTTCAATGACACATCCCAAGAACTTTGGCTGATGCAATCTCTGAGAG 1380  
DB 1473 CCGGAGAAAGTTCAATGACACATCCCAAGAACTTTGGCTGATGCAATCTCTGAGAG 1532  
QY 1381 GAAAGACATGTGCTGAGTGGCTCTTATTAATCTACTGACTAAGAAATGAAGAACTTAA 1440  
DB 1533 GAAAGACATGTGCTGAGTGGCTCTTATTAATCTACTGACTAAGAAATGAAGAACTTAA 1592  
QY 1441 GAGCTGTGAGAGCGAGCTATGCGCGCGGCGAAGACCGAGCAACCAAGAGAGA 1500  
DB 1593 GAGCTGTGAGAGCGAGCTATGCGCGCGGCGAAGACCGAGCAACCAAGAGAGA 1552  
QY 1501 GCAGCAGACAGCAGCAGCAGCAGCAGAGCCCATGCCCGGACAGCGCAGAGAGAGAA 1560  
DB 1653 GCAGCAGACAGCAGCAGCAGCAGCAGCAGAGCCCATGCCCGGACAGCGCAGAGAGAGAA 1712  
QY 1561 AGATGAGAGAGAGAGAGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
DB 1713 AGATGAGAGAGAGAGAGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1772  
QY 1621 CGACAAAGAGAGCTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
DB 1773 CGACAAAGAGAGCTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1832  
QY 1681 GAAAGAGCTGTGCGCTTCCAAAGCGCGCAAACTGCGCAACGCGAGAGAGAGAGAGAG 1740

Db 1833 GAAGAGGCTGTGGCTTCCAAAGCCGCAAACTGCTCAACAGCCAGGAGAGACGCAAAAG 1892  
QY 1741 CCGCATCAACCCGCTCAATGGCTAATAGAGCCAAACAGCGAGAGGACATCAACCCCCAGCA 1800  
Db 1893 CCGCATCAACCCGCTCAATGGCTAATAGAGCCAAACAGCGAGAGGACATCAACCCCCAGCA 1952  
QY 1801 GAGCCCGGAGCTGGCCCTCCATGAGCTGAATGAGAGTTCTCGCTGACAGAAAGAAAT 1860  
Db 1953 GAGCCCGGAGCTGGCCCTCCATGAGCTGAATGAGAGTTCTCGCTGACAGAAAGAAAT 2012  
QY 1861 GGAACACGCAAGAAAGGTCTCTGGAACAAGCCGCAACCTGTGGCCCATGCCCCGAT 1920  
Db 2013 GGAACACGCAAGAAAGGTCTCTGGAACAAGCCGCAACCTGTGGCCCATGCCCCGAT 2072  
QY 1921 GGTGGGCTTCAAGACTGTGTGCACTGTAAAGTCTTCACTTCAACTAAGAAAGGCA 1980  
Db 2073 GGTGGGCTTCAAGACTGTGTGCACTGTAAAGTCTTCACTTCAACTAAGAAAGGCA 2132  
QY 1981 GAACTTCATGATGATCTTTCAGCAGCAACAGCTGAAGATGAGAAAGAGAGAGCGCG 2040  
Db 2133 GAACTTCATGATGATCTTTCAGCAGCAACAGCTGAAGATGAGAAAGAGAGAGCGCG 2192  
QY 2041 GAGGAAGAAAGAAAGAGCGCGCGCGCGCGCGAGGAGGCTGATTCGCCCGCTGGT 2100  
Db 2193 GAGGAAGAAAGAAAGAGCGCGCGCGCGCGCGAGGAGGCTGATTCGCCCGCTGGT 2252  
QY 2101 GAGGATGAGAGATGAGAGCTGCGGCGCTGACGCGAAATGAGAGAGATGTTGAGAGA 2160  
Db 2253 GAGGATGAGAGAGATGAGAGCGCTCGGGCGTGAAGGGAATGAGAGAGAGATGTTGAGAGA 2312  
QY 2161 GGCTGAAGCTTATCATGCTCTTGGAATGAGGTGCCAGAGGGGAATGAGATGGCCCAAC 2220  
Db 2313 GGCTGA-----AGC 2321  
QY 2221 CACTCTCAACAACAGCTCAGACACCGAGAGCATCCCTCTCTCTCAACTGAGAGCGGCA 2280  
Db 2332 CACTCTCAACAACAGCTCAGACACCGAGAGCATCCCTCTCTCTCAACTGAGAGCGGCA 2381  
QY 2281 GGAACAAGGGCGAATGGGCCCAAGCCCAACCCCAACCCCTGGGCGCGAGCGGCCAACC 2441  
Db 2382 GGAACAAGGGCGAATGGGCCCAAGCCCAACCCCAACCCCTGGGCGCGAGCGGCCAACC 2441  
QY 2341 AGGCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCC 2400  
Db 2442 AGGCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCC 2501  
QY 2401 TGAAGCCACCGGAGCGCCCTCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAAC 2460  
Db 2502 TGAAGCCACCTTAAGCCCTCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAAC 2561  
QY 2461 TGTGTTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
Db 2552 TGTGTTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2621  
QY 2551 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
Db 2622 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2681  
QY 2591 GCGCCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
Db 2682 GCGCCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2741  
QY 2641 GAGAGCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
Db 2742 GAGAGCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2801  
QY 2701 GCGCAGGGCCACCACTGCGCAAGAGCTCGGGCGCCCGCCAGAGCAAGGACTTCAAGTGTAC 2760  
Db 2802 GCGCAGGGCCACCACTGCGCAAGAGCTCGGGCGCCCGCCAGAGCAAGGACTTCAAGTGTAC 2861  
QY 2761 CTGCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
Db 2862 CTGCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2921

QY 2821 AAGGCCAGCCTCTCAACCCCGAGCTGGGAGACCCCGGGCAATGCTTCAACCCAGAGACC 2880  
Db 2922 AAGGCCAGCCTCTCAACCCCGAGCTGGGAGACCCCGGGCAATGCTTCAACCCAGAGACC 2981  
QY 2881 ACTGACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
Db 2982 ACTGACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3041  
QY 2941 AGTCCATAGACCCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000  
Db 3042 AGTCCATAGACCCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3101  
QY 3001 GCGACCGCAAAACCTGACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
Db 3102 GCGACCGCAAAACCTGACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3161  
QY 3061 GGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3100  
Db 3162 GGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3221  
QY 3101 ----GCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3156  
Db 3222 CCGAGCCTTCCGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3281  
QY 3157 CCTGCGCTTCCCGTGGCTGCGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3216  
Db 3282 CTTGCGCTTCCCGTGGCTGCGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3341  
QY 3217 CTGAGCCTTCTCTCAAGCTTCAACCTGCTCAACCTGCTGAGAGAGAGAGAGAGAGAG 3276  
Db 3342 CTGAGCCTTCTCTCAAGCTTCAACCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3401  
QY 3277 CCGGCGCGTCTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3336  
Db 3402 CCGGCGCGTCTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3461  
QY 3337 CAAGCAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3396  
Db 3462 CAAGCAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3521  
QY 3397 GCTCCAGTCTCCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3456  
Db 3522 GCTCCAGTCTCCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3581  
QY 3457 GCGCCCTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3516  
Db 3582 GCGCCCTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3641  
QY 3517 GTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3576  
Db 3642 GTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3701  
QY 3577 GTCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3636  
Db 3702 GTCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3761  
QY 3637 TCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3696  
Db 3762 TCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3821  
QY 3697 CAGCCAGCTGAGCTCTGTAACAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3756  
Db 3822 CAGCCAGCTGAGCTCTGTAACAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3881  
QY 3757 GAGTGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3816  
Db 3882 GAGTGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3941  
QY 3817 CAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3876  
Db 3942 CAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4001



OY	3877	GGACGGCGAAGAGCAGCTACAGAACCCCGCCCAATGAGACGGCGCCGCCCAAGCGCACTATGAA	3936
Db	4002	GGACGGCGAAGAGCAGCTACAGAACCCCGCCCAATGAGACGGCGCCGCCCAAGCGCACTATGAA	4061
OY	3937	CATGATGGAAGGAGCGCGCTGGGGCAGAGCCATCTCTCAGCCAGCATTCGAAGGATCTCATGGG	3996
Db	4062	CATGATGGAAGGAGCGCGCTGGGGCAGAGCCATCTCTCAGCCAGCATTCGAAGGATCTCATGGG	4121
OY	3997	CGGTGCATCCCGCGGAGCGACACAGCCGCCACCACTCAAGAGCAGCACCATTCG	4056
Db	4122	CGGTGCATCCCGCGGAGCGACACAGCCGCCACCACTCAAGAGCAGCACCATTCG	4181
OY	4057	CGGGTTCATCAACAAGGAGATCCCTTCGGTCTTAAGTGAAGGCACAGAGAGACTACCTGCG	4116
Db	4182	CGGGTTCATCAACAAGGAGATCCCTTCGGTCTTAAGTGAAGGCACAGAGAGACTACCTGCG	4241
OY	4117	TCGGGAGGCGAAGCTCTTAAGACGGGAGAGAGCGCTCCGGCCGCCACCGCCCTCAAGGGAA	4176
Db	4242	TCGGGAGGCGAAGCTCTTAAGACGGGAGAGAGCGCTCCGGCCGCCACCGCCCTCAAGGGAA	4301
OY	4177	CCTGACCGAGGCGTACAAAGACGACAGGCCCTTGGGCCCTCGAAGCTGAAGCGGCGCCATGA	4236
Db	4302	CCTGACCGAGGCGTACAAAGACGACAGGCCCTTGGGCCCTCGAAGCTGAAGCGGCGCCATGA	4361
OY	4237	GGGCGCTGGTGGCCACAGCTGTAAGAGAGGGCGGGCCGCTCATCTCATGAGATCCCGCGAGGA	4296
Db	4362	GGGCGCTGGTGGCCACAGCTGTAAGAGAGGGCGGGCCGCTCATCTCATGAGATCCCGCGAGGA	4421
OY	4287	GCTGGGGCAACGCCCGAGCTGCCCTTGGCCCGCGCGCGCTCAAGAGAGGGCTCATCAC	4356
Db	4422	GCTGGGGCAACGCCCGAGCTGCCCTTGGCCCGCGCGCGCTCAAGAGAGGGCTCATCAC	4481
OY	4357	GCAGGGGCAACCCCGCTCAAGTACGACACCGGGCGGTCACACATGGGCTCAAAAAGACGAA	4416
Db	4482	GCAGGGGCAACCCCGCTCAAGTACGACACCGGGCGGTCACACATGGGCTCAAAAAGACGAA	4541
OY	4417	CGTACGCTCCCTCATTCGCGACGCCCGCGCGAGCGTTCCACCCGCTGACACCGCTGATGT	4476
Db	4542	CGTACGCTCCCTCATTCGCGACGCCCGCGCGAGCGTTCCACCCGCTGACACCGCTGATGT	4601
OY	4477	GATGGCGGACGCCCGGGGACCTGGAACGTGGCCCTGACGAGGAGAGCCCTGAAGACCGGGCC	4536
Db	4602	GATGGCGGACGCCCGGGGACCTGGAACGTGGCCCTGACGAGGAGAGCCCTGAAGACCGGGCC	4661
OY	4537	AGGGACCGCCACAGCAGCTCGGGGGGCTCCATTGGCGCGCGGCCGCCCGGCTCATTTGCGCTGA	4596
Db	4662	AGGGACCGCCACAGCAGCTCGGGGGGCTCCATTGGCGCGCGGCCGCCCGGCTCATTTGCGCTGA	4721
OY	4597	GCTGGGTAAAGCCGCGGACAGACCCCGCTGACCTATGAGGACCAAGGGGCAACCTTTGCGCG	4656
Db	4722	GCTGGGTAAAGCCGCGGACAGACCCCGCTGACCTATGAGGACCAAGGGGCAACCTTTGCGCG	4781
OY	4657	CCAACCTCCCAAGAGTTTGGCCCTGACCATGCGGGAGGCCACCGCGCGCTGACAGAGAGG	4716
Db	4782	CCAACCTCCCAAGAGTTTGGCCCTGACCATGCGGGAGGCCACCGCGCGCTGACAGAGAGG	4841
OY	4717	CAGCCTTTGCTCAAGCAAGGACCTCCAGAGACGGAAGCTGAAGCTCGACGCGCTCGTGAGAT	4776
Db	4842	CAGCCTTTGCTCAAGCAAGGACCTCCAGAGACGGAAGCTGAAGCTCGACGCGCTCGTGAGAT	4901
OY	4777	CGCCAAAGTCCCGGACACAGACCTGTGCCCGAGACACCAACCCCAATCTTGCCCTATGAA	4836
Db	4902	CGCCAAAGTCCCGGACACAGACCTGTGCCCGAGACACCAACCCCAATCTTGCCCTATGAA	4961
OY	4837	GCACTGCTTTGGGGGCTGAGTGGCGTGGACCTGTATTTGACAGCCACATCCCGCTGGGCTT	4896
Db	4962	GCACTGCTTTGGGGGCTGAGTGGCGTGGACCTGTATTTGACAGCCACATCCCGCTGGGCTT	5021
OY	4897	GGAACCCCACTTCATACCCCGGGGACATCCCTTGGAGCGACCGGCTGCTATCTACCTGGC	4956
Db	5022	GGAACCCCACTTCATACCCCGGGGACATCCCTTGGAGCGACCGGCTGCTATCTACCTGGC	5081
OY	4957	CGGACACTGGGCCCCCAACCCCACTTACCGGACCTGTATCCCAACCTTACCTCATCCGGG	5016

Db	5082	CCGACACCTGAGCCCCCCAAACCCACCTACCCGACCTGACCCACCCCTACCTCATCCGGGG	5141
Oy	5017	CTACCCCGACACGGCGCGCTGAGAACCGGAGACCATCATCAATGACTACATCACTC	5076
Db	5142	CTACCCCGACACGGCGCGCTGAGAACCGGAGACCATCATCAATGACTACATCACTC	5201
Oy	5077	GCAGCAGATGACCAACAACACGGCCACCGCCATGCGCCACGAGAGTGATATGCTGAGGGG	5136
Db	5202	GCAGCAGATGACCAACAACACGGCCACCGCCATGCGCCACGAGAGTGATATGCTGAGGGG	5261
Oy	5137	CTCTCGACCCCGGAGTCTCTCGTGAGCACTCAACTACGCTGGGGGTCCCGGAGCATCAT	5196
Db	5262	CTCTCGACCCCGGAGTCTCTCGTGAGCACTCAACTACGCTGGGGGTCCCGGAGCATCAT	5321
Oy	5197	CGACTGTCCCAAGTGCACACCTGCTGTGCTGTGCCCCCGACACGAGGACCCGACG	5256
Db	5322	CGACTGTCCCAAGTGCACACCTGCTGTGCTGTGCCCCCGACACGAGGACCCGACG	5381
Oy	5257	CACCGCCATGACCGCGCTTGGCTTACTCTCCACCGCGGCCAGCGCTTACAGACGCGCA	5316
Db	5382	CACCGCCATGACCGCGCTTGGCTTACTCTCCACCGCGGCCAGCGCTTACAGACGCGCA	5441
Oy	5317	CACGAGTCTCCCACTCTCCCGAGAGTCCACAACAATTGACAAACCAACACACGTC	5376
Db	5442	CACGAGTCTCCCACTCTCCCGAGAGTCCACAACAATTGACAAACCAACACACGTC	5501
Oy	5377	CTCGTCCGAGCGGAGCGAGACCGGAGTGAAGCGGACCGGGATTCGGAGCGGAAAA	5436
Db	5502	CTCGTCCGAGCGGAGCGAGACCGGAGTGAAGCGGACCGGGATTCGGAGCGGAAAA	5561
Oy	5437	GTCATCTCTCAAGTCCACCAACGACGGTGAAGACCGACCATCTTGAAGACCTTGATCAGA	5496
Db	5562	GTCATCTCTCAAGTCCACCAACGACGGTGAAGACCGACCATCTTGAAGACCTTGATCAGA	5621
Oy	5497	GCAGAGAGGGGACGACAGGCGGACGACGCGCGGGGGTGGGGGCAACAAGCCGCGCGC	5556
Db	5622	GCAGAGAGGGGACGACAGGCGGACGACGCGCGGGGGTGGGGGCAACAAGCCGCGCGC	5681
Oy	5557	CTCCCACTCCCATGACCACACGACACTGCGCCACTCTCCCTCGGACCCAGATGCGCTCCA	5616
Db	5682	CTCCCACTCCCATGACCACACGACACTGCGCCACTCTCCCTCGGACCCAGATGCGCTCCA	5741
Oy	5617	GCAGAGACCCAGTGTGCTTCAACAACAAGGATGAAGGGTATCATCACCGCTGTGAAGCC	5676
Db	5742	GCAGAGACCCAGTGTGCTTCAACAACAAGGATGAAGGGTATCATCACCGCTGTGAAGCC	5801
Oy	5677	CAGCAGCCCAACGGTCTCTGAGGTCCACTTCACCTTCTCAACCCGTTTCGCCAGCTGCCAC	5736
Db	5802	CAGCAGCCCAACGGTCTCTGAGGTCCACTTCACCTTCTCAACCCGTTTCGCCAGCTGCCAC	5861
Oy	5737	ATTCCCACTCCGACACCAATGACCACCTGAGGGAGGACCCCTGATGGAGGTCTACACCT	5796
Db	5862	ATTCCCACTCCGACACCAATGACCACCTGAGGGAGGACCCCTGATGGAGGTCTACACCT	5921
Oy	5797	CATGAGACCGGTCTTCTGCTGCCAAGAGGACCCCGGGTTCGCGCCGACGAGACCGGCCG	5856
Db	5922	CATGAGACCGGTCTTCTGCTGCCAAGAGGACCCCGGGTTCGCGCGCCGACGAGACCGGCCG	5981
Oy	5857	AGCAGACACCGGCATGCTTCTGCTGCCAAGCCCCCAAGCCCGCTCCGGGCTGAGACCCGC	5916
Db	5982	AGCAGACACCGGCATGCTTCTGCTGCCAAGCCCCCAAGCCCGCTCCGGGCTGAGACCCGC	6041
Oy	5917	CTCCTCCCGACGACAAAGGCTCGAGGCCCGGAGCCCTGAGTCTCTGCTCTGAGCAGC	5976
Db	6042	CTCCTCCCGACGACAAAGGCTCGAGGCCCGGAGCCCTGAGTCTCTGCTCTGAGCAGC	6101
Oy	5977	CACCATGACCCGACCCCTGCGAAGAACCTTCGACCTCAACACCGCAGCCCGGACCGCGC	6036
Db	6102	CACCATGACCCGACCCCTGCGAAGAACCTTCGACCTCAACACCGCAGCCCGGACCGCGC	6161
Oy	6037	GGCGCACCTGCTCGGCTTCGAGCCCGGACCTCGGAAAAAGTCAAAAGTAAACCTTTTC	6096

Db 6162 GAGCCACCTGCTGGGCTCGAGCCCGACCCGGAAAAAGACTCAAGTAAACCTTTTC 6221  
QY 6097 CATCCAGGAATCGAATCTCCGTTCTTGGGTTACACCGCAGACGTAACGCCCGGAAG 6156  
Db 6222 CATCCAGGAATCGAATCTCCGTTCTTGGGTTACACCGCAGACGTAACGCCCGGAAG 6281  
QY 6157 GGTGAAGCCCGTCAAGCCCTGTGAGCTCACCCAGTCTGACCCAGACAGAGGGGCTCCCAA 6216  
Db 6282 GGTGAAGCCCGTCAAGCCCTGTGAGCTCACCCAGTCTGACCCAGACAGAGGGGCTCCCAA 6341  
QY 6217 GCACCTGGAAAGGCTCGACAAAGAGCCACTGGAGGGGGAGCTGGGGCCCAACAGCCAGG 6276  
Db 6342 GCACCTGGAAAGGCTCGACAAAGAGCCACTGGAGGGGGAGCTGGGGCCCAACAGCCAGG 6401  
QY 6277 CCCCGTGAAGCTTGGCGGGAGGCGGCCCACTCCCAACCTGCGCGCTGAGAG 6336  
Db 6402 CCCCGTGAAGCTTGGCGGGAGGCGGCCCACTCCCAACCTGCGGGCGCTGAGAG 6461  
QY 6337 CCAAGCCCTGTCAGCCCGCTGCTCCAGACCGCCCGAGGGGTCAAAAGTCAACAGCGGT 6396  
Db 6462 CCAAGCCCTGTCAGCCCGCTGCTCCAGACCGCCCGAGGGGTCAAAAGTCAACAGCGGT 6521  
QY 6397 GGTCAACCTGGCCAGACATCATGAGGGTCAACAAGAGCTACACCGGCAACACCC 6456  
Db 6522 GGTCAACCTGGCCAGACATCATGAGGGTCAACAAGAGCTACACCGGCAACACCC 6581  
QY 6457 ACAGAGCTCAGACGACCCCTGCGCCGCCCTTACTCTTCCTCGGGGCGAGCTGCC 6516  
Db 6582 ACAGAGCTCAGACGACCCCTGCGCCGCCCTTACTCTTCCTCGGGGCGAGCTGCC 6641  
QY 6517 GGTCTGGAACCTTCCGCGCGCCACCCAGTGAACCTTGAACCTCCGCCCGAGCATGTGC 6576  
Db 6642 GGTCTGGAACCTTCCGCGCGCCACCCAGTGAACCTTGAACCTCCGCCCGAGCATGTGC 6701  
QY 6577 CCGGCGCGGTGCTCCCCACAGCGAAGGGGGAGAGGTCTCAAGGCCAAAGAGC 6636  
Db 6702 CCGGCGCGGTGCTCCCCACAGCGAAGGGGGAGAGGTCTCAAGGCCAAAGAGC 6761  
QY 6637 GTCCGTCTTGGGTGTTGTTGAGACGGTATTTGAACTGTGTCCCAACCGAGAGCATGAC 6696  
Db 6762 GTCCGTCTTGGGTGTTGTTGAGACGGTATTTGAACTGTGTCCCAACCGAGAGCATGAC 6821  
QY 6697 GAGGCCAGGGGCTCTCCCGAGTGTGTGTATCCCGCTGCTTACCGGAGTGGGAAACAGC 6756  
Db 6822 GAGGCCAGGGGCTCTCCCGAGTGTGTGTATCCCGCTGCTTACCGGAGTGGGAAACAGC 6881  
QY 6757 GAGGCCAGGAGTGGGCTCCAGTCTCAAGGCAACACAGCCAGCCGCAACCTTCTT 6816  
Db 6882 GAGGCCAGGAGTGGGCTCCAGTCTTCCAGGCAACACAGCCAGCCGCAACCTTCTT 6941  
QY 6817 CAGCAAGCTGACCCGAGACAACTCCGCCATGTTCAAGTCCAAAGACAAAGAGATCAACA 6876  
Db 6942 CAGCAAGCTGACCCGAGACAACTCCGCCATGTTCAAGTCCAAAGACAAAGAGATCAACA 7001  
QY 6877 GAAGCTGAACACCCAAACCCGAGATGAGCTGAATATCAATATCAGCCAGCTGGAGCGA 6936  
Db 7002 GAAGCTGAACACCCAAACCCGAGATGAGCTGAATATCAATATCAGCCAGCTGGAGCGA 7061  
QY 6937 GATCTTCAATATGCCCCGATCAACCGGAACAGGCTTATGACCTTATGAAGCAGCGGT 6996  
Db 7062 GATCTTCAATATGCCCCGATCAACCGGAACAGGCTTATGACCTTATGAAGCAGCGGT 7121  
QY 6997 GCAAGAACATGCCAGACCAACATGAGGCTGAGAGGCAATTAATTAGAAAGGCACTCATGG 7056  
Db 7122 GCAAGAACATGCCAGACCAACATGAGGCTGAGAGGCAATTAATTAGAAAGGCACTCATGG 7181  
QY 7057 TAAATATGACAGTGGGAAAGAGTCCCGCGCTCAGCGCAATGCTTTTAAACCTCTGA 7116  
Db 7182 TAAATATGACAGTGGGAAAGAGTCCCGCGCTCAGCGCAATGCTTTTAAACCTCTGA 7241  
QY 7117 TGCCAGTGCAGGCTGCGCGCTGCTATGCCCCATACCGCTGTGACGAGCGAGTGAACA 7176  
Db 7242 TGCCAGTGCAGGCTGCGCGCTGCTATGCCCCATACCGCTGTGACGAGCGAGTGAACA 7301

QY 7177 CACACTCACTTGAGGATGGGCGGAGAGGCCAAAGTCTTGTGACAGACCCAGGACCG 7236  
Db 7302 CACACTCACTTGAGGATGGGCGGAGAGGCCAAAGTCTTGTGACAGACCCAGGACCG 7361  
QY 7237 AAAAGCCAAAGTCCCCGGGCGCGGCTGCAATCTGGGGAACCGGCCAACCTCTGTCTCTC 7296  
Db 7362 AAAAGCCAAAGTCCCCGGGCGCGGCTGCAATCTGGGGAACCGGCCAACCTCTGTCTCTC 7421  
QY 7297 AGTGACTCGAGGAGACTCAACCGCCGAGACCGCGCTACCAACCGCGTGTGGAGGA 7356  
Db 7422 AGTGACTCGAGGAGACTCAACCGCCGAGACCGCGCTACCAACCGCGTGTGGAGGA 7481  
QY 7357 CAGGCTCTGTCCGAGGTTCCAGCGCATTCCTTACAAACCCCTGATCATGCGCTGA 7416  
Db 7482 CAGGCTCTGTCCGAGGTTCCAGCGCATTCCTTACAAACCCCTGATCATGCGCTGA 7541  
QY 7417 GCGGGGTGATATGATTTCCCAACCCGAGCTCCCCCGGGCAGCGGGCCCTTGC 7476  
Db 7542 GCGGGGTGATATGATTTCCCAACCCGAGCTCCCCCGGGCAGCGGGCCCTTGC 7601  
QY 7477 TGAGCCCCACACGCTGGAGAGAGAGCCCAAGGCACTGCTGTGCTGAGTACGAGAC 7536  
Db 7602 TGAGCCCCACACGCTGGAGAGAGAGCCCAAGGCACTGCTGTGCTGAGTACGAGAC 7661  
QY 7537 ACTCTCCGACAGCGAGTACTCAGAACAGGCGGGGGGGGGCGGGCGGTGEBRSPRLUG 7596  
Db 7662 ACTCTCCGACAGCGAGTACTCAGAACAGGCGGGGGGGGGCGGGCGG-----TG 7711  
QY 7597 TCAAGTCCCAAGAGCCACAGGAAGCGCTGTGAGAGACGGGGCGGCTGCCGACCTCCCC 7656  
Db 7712 TCAAGTCCCAAGAGCCACAGGAAGCGCTGTGAGAGACGGGGCGGCTGCCGACCTCCCC 7771  
QY 7657 AACCAAGGAAGAGCCCTGAGTCCGCTGCGCCCTCAATCCATCTGTCCAGAGCG 7716  
Db 7772 AACCAAGGAAGAGCCCTGAGTCCGCTGCGCCCTCAATCCATCTGTCCAGAGCG 7831  
QY 7717 GCATCTTTCCTGTCTAAAGCTTAACTTAAGATTCGCCGCCCGGCTGAGCTGTGACA 7776  
Db 7832 GCATCTTTCCTGTCTAAAGCTTAACTTAAGATTCGCCGCCCGGCTGAGCTGTGACA 7891  
QY 7777 CCTTACTAGGAGATGTTTACTGTGTGTGCGGAAGGAGGGGAGGGCCGGGAGAGGG 7836  
Db 7892 CCTTACTAGGAGATGTTTACTGTGTGTGCGGAAGGAGGGGAGGGCCGGGAGAGGG 7951  
QY 7837 GCAAGCGAGCGGTGGCAGGCAACACAGCGGCGCAAGGCGGCCAAGGAGCCCAAGCAG 7896  
Db 7952 GCAAGCGAGCGGTGGCAGGCAACACAGCGGCGCAAGGCGGCCAAGGAGCCCAAGCAG 8011  
QY 7957 GATGACACAGCACTCCAGGCACTGCTCCCGGATGCAATTTGGAACCAAGTCTTAA 7956  
Db 8012 GATGACACAGCACTCCAGGCACTGCTCCCGGATGCAATTTGGAACCAAGTCTTAA 8071  
QY 7957 CTGAGCTGCGAGCCCGCGCGCTCTCCCTCCGCTCCCATCCCGCTTACCGCTGTGAGCAG 8016  
Db 8072 CTGAGCTGCGAGCCCGCGCGCTCTCCCTCCGCTCCCATCCCGCTTACCGCTGTGAGCAG 8131  
QY 8017 ATGAGCGAGGCGCTGTCAGCGCCGAGTGGCGTGTTCGGGTCCCGCAGACTGCCCA 8076  
Db 8132 ATGAGCGAGGCGCTGTCAGCGCCGAGTGGCGTGTTCGGGTCCCGCAGACTGCCCA 8191  
QY 8077 GCCAACGAGATTTGCTGGAACCAAGTCAAGCCAGGTGGCGGACAAAGGGCCAGGTGCG 8136  
Db 8192 GCCAACGAGATTTGCTGGAACCAAGTCAAGCCAGGTGGCGGACAAAGGGCCAGGTGCG 8251  
QY 8137 GCTTGGGGGGAACGAGTCTCCAGAGACTGATGTTTTTTTCAACATCGTTGCCGAG 8196  
Db 8252 GCTTGGGGGGAACGAGTCTCCAGAGACTGATGTTTTTTTCAACATCGTTGCCGAG 8311  
QY 8197 CGGTGGGAAGGAAGGCAAGTAAATGATGTTGTTTCAAGGATATTTTTGATAC 8256  
Db 8312 CGGTGGGAAGGAAGGCAAGTAAATGATGTTGTTTCAAGGATATTTTTGATAC 8371

QY 8257 CTTCATGATTAATTCAGATGTTTTACGAGAGAGACTTACCCAGTATTACTCTGC 8316  
XX |||||  
DB 8372 CTTCATGATTAATTCAGATGTTTTACGAGAGAGACTTACCCAGTATTACTCTGC 8431  
QY 8317 TGTGCTTTGATCTGCTTACCGTTCAAGAGCGTGTGACGCGACAGTCCGTGACCC 8376  
DB 8432 TGTGCTTTGATCTGCTTACCGTTCAAGAGCGTGTGACGCGACAGTCCGTGACCC 8491  
QY 8377 CATCATCTCCAGAGACCAAGGGGCGGGGACTGCTGCTACCGCCGCTGTCTCCCTC 8436  
DB 8492 CATCATCTCCAGAGACCAAGGGGCGGGGACTGCTGCTACCGCCGCTGTCTCCCTC 8551  
QY 8437 CCTCCCTTCTTGAGGAGATGATTCAGATGCGTATTCTGTGGCCGCCATTTGCGAGGG 8496  
DB 8552 CTTCCCTTCTTGAGGAGATGATTCAGATGCGTATTCTGTGGCCGCCATTTGCGAGGG 8611  
QY 8497 TGTGCTTTGATCTGCTTACCATTTACACACGCTGTTCTAATTTAAAGCGAATTACTCCAAA 8556  
DB 8612 TGTGCTTTGATCTGCTTACCATTTACACACGCTGTTCTAATTTAAAGCGAATTACTCCAAA 8671  
QY 8557 AAAAAAAAAAAAAA 8571  
DB 8672 AAAAAAAAAAAAAA 8686  
RESULT 10  
ID ADQ18920  
ADQ18920 standard; DNA; 8686 BP.  
XX  
AC ADQ18920;  
XX  
DT 26-AUG-2004 (first entry)  
XX  
DS Human soft tissue sarcoma-upregulated DNA - SEQ ID 1739.  
XX  
KW soft tissue sarcoma; cytosstatic; gene therapy; vaccine; screening; human;  
XX db.  
OS Homo sapiens.  
XX  
PN WO2004048938-A2.  
XX  
PD 10-JUN-2004.  
XX  
PR 26-NOV-2003; 2003WO-US038193.  
XX  
PR 26-NOV-2002; 2002US-0429739P.  
XX  
PA (PROT-) PROTEIN DESIGN LABS INC.  
XX  
PI Aziz N, Ginsburg WM, Zlotnik A;  
XX  
XX WPI; 2004-441208/41.  
DR  
XX  
XX  
PT Early detection of soft tissue sarcoma comprises determining expression  
PT of a gene in a first soft tissue sample and a normal soft tissue sample  
PT and comparing the gene expression, also useful in treating soft tissue  
PT sarcoma.  
PS  
XX  
XX Example 2; SEQ ID NO 1739; 210pp; English.  
CC The invention relates to a novel method for detecting soft tissue sarcoma  
CC which comprises obtaining a first soft tissue sample from an individual  
CC and a normal soft tissue sample from the same or different individual,  
CC determining the expression of a gene in both samples and comparing the  
CC expression of the gene in both soft tissue samples, where a higher level  
CC of protein expression in the first soft tissue sample indicates the  
CC presence of soft tissue sarcoma. The method of the invention has  
CC cytostatic applications and may be useful for detecting soft tissue  
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic  
CC acid sequences may be useful in diagnostic and screening applications.  
CC The current sequence is that of a human soft tissue sarcoma-upregulated  
CC DNA of the invention. The current sequence is not shown within the

CC specification per se but was submitted in CD format by the inventor.  
XX  
SQ Sequence 8686 BP; 1881 A; 3069 C; 2562 G; 1174 T; 0 U; 0 Other;  
Query Match 97.2%; Score 8324.6; DB 12; Length 8686;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 8472; Conservative 1; Mismatches 34; Indels 88; Gaps 4;  
QY 1 CATGTGGGCTCCACACAGCTTGTGACACAGAGGTGAGGGCCACTGAGCCCGCTACCC 60  
DB 156 CATGTGGGCTCCACACAGCTTGTGACACAGAGGTGAGGGCCACTGAGCCCGCTACCC 215  
QY 61 GCCCCACAGCCTTCTTCTTACCCAGTGCAGATCGCCCGGACGACACCGAGCGTGGGCTCT 120  
DB 216 GCCCCACAGCCTTCTTCTTACCCAGTGCAGATCGCCCGGACGACACCGAGCGTGGGCTCT 275  
QY 121 GGAATGACAGACCACTCCCGGACTATGCTTCCACTGTGCGCGGCTCCATCTCCA 180  
DB 276 GGAATGACAGACCACTCCCGGACTATGCTTCCACTGTGCGCGGCTCCATCTCCA 335  
QY 181 GCCCCAGCGGCGGAGGCGCTCCCTGCTGTGAGTTCCAGCCCGGGAATGAAACGGTCCCA 240  
DB 336 GCCCCAGCGGCGGAGGCGCTCCCTGCTGTGAGTTCCAGCCCGGGAATGAAACGGTCCCA 395  
QY 241 GGAATGACAGCTTGGGCGGAGTCCCACTATACCTGCCGAGCTGGGGAATGACAGAT 300  
DB 396 GGAATGACAGCTTGGGCGGAGTCCCACTATACCTGCCGAGCTGGGGAATGACAGAT 455  
QY 301 GGAATGATTTGAAAGCAAGCGGCTTGGGCTGAGAGCTGTGCTGACCCCTGTGCGACC 360  
DB 456 GGAATGATTTGAAAGCAAGCGGCTTGGGCTGAGAGCTGTGCTGACCCCTGTGCGACC 515  
QY 361 GTTACCCCTGTGCGGCGGAGCGGCGGAGGCTGAGAGCTGAGAGCTGAGAGCGGTAG 420  
DB 516 GTTACCCCTGTGCGGCGGAGCGGCGGAGGCTGAGAGCTGAGAGCTGAGAGCGGTAG 575  
QY 421 CCTGACGGGCGAAGTGGAAACCGGTGTCTCCCGGAGCGGCGGAGCTGAGAGCTGAGAG 480  
DB 576 CCTGACGGGCGAAGTGGAAACCGGTGTCTCCCGGAGCGGCGGAGCTGAGAGCTGAGAG 635  
QY 481 GGAAGCTGTGCGGCGGAGCGGCTGCTCCAGGAGAGAGCTGATCCAGACATGAGCCGCTGAG 540  
DB 636 GGAAGCTGTGCGGCGGAGCGGCTGCTCCAGGAGAGAGCTGATCCAGACATGAGCCGCTGAG 695  
QY 541 CCGAGAGATCAACATGCTGAGAGAGAGAGCTGATCCAGACATGAGCCGCTGAGAG 600  
DB 696 CCGAGAGATCAACATGCTGAGAGAGAGAGCTGATCCAGACATGAGCCGCTGAGAG 755  
QY 601 GGAAGAGAGAGCTGCGCAAGCGGCTGCTCCAGGAGAGAGAGCTGAGAGAGCTGAGAG 660  
DB 756 GGAAGAGAGAGCTGCGCAAGCGGCTGCTCCAGGAGAGAGAGCTGAGAGAGCTGAGAG 815  
QY 661 GTTGAAGACCGGAGCGGCTGCTGCAATCTACGAGAGAGAGAGAGAGAGCTGAGAG 720  
DB 816 GTTGAAGACCGGAGCGGCTGCTGCAATCTACGAGAGAGAGAGAGAGAGCTGAGAG 875  
QY 721 TGCACATGCGGATTTGAGAGAGGCTGCGGCGGCGGAGAGAGAGCTGAGAGAGCTGAGAG 780  
DB 876 TGCACATGCGGATTTGAGAGAGGCTGCGGCGGCGGAGAGAGAGCTGAGAGAGCTGAGAG 935  
QY 781 CTCGAGACCGGAGCGGAGTATCATGAGAACTCAAAATTAACGAGAGAGAGAGAGT 840  
DB 936 CTCGAGACCGGAGCGGAGTATCATGAGAACTCAAAATTAACGAGAGAGAGAGAGT 995  
QY 841 AATCTTGTACTTCAAGAGAGAGAGATCAAGCTGAGAGAGAGAGAGAGAGAGTCTGCA 900  
DB 996 AATCTTGTACTTCAAGAGAGAGAGATCAAGCTGAGAGAGAGAGAGAGAGAGTCTGCA 1055  
QY 901 GGGCTATGACCGAGCTCATGAGAGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
DB 1056 GGGCTATGACCGAGCTCATGAGAGAGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1115  
QY 961 GGGCGGCGGGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020

Db 1116 CCGGGGGGGGCGCAAGAGAGAGATTGCGAGATACGAGAAAGCAGTTCCCTGAGAT 1175  
Qy 1021 CCGCAAGCAGCGCGAGCTGCGAGAGCGATGCAAGCAGGAGGCGCGGCGAGTGG 1080  
Db 1176 CCGCAAGCAGCGCGAGCTGCGAGAGCGATGCAAG--AGGATGGCGAGGCGGCGAGTGG 1232  
Qy 1081 GCTGTCAATGTGCGCGCGCGCGAGAGCAGAGGTGTCAAGATCATCGATGAGCTCTC 1140  
Db 1233 GCTGTCAATGTGCGCGCGCGCGAGAGCAGAGGTGTCAAGATCATCGATGAGCTCTC 1292  
Qy 1141 AGAGCAGAGAAACCTGAGAGACAGATGCGCGAGCTGCGCTGATTCGCGCCATGCTGTA 1200  
Db 1293 AGAGCAGAGAAACCTGAGAGACAGATGCGCGAGCTGCGCTGATTCGCGCCATGCTGTA 1352  
Qy 1201 CGACGCTGACGAGCGCGCATGATGATCAACATGAAAGGAGCTTATGAGCGAGCCCAT 1260  
Db 1353 CGACGCTGACGAGCGCGCATGATGATCAACATGAAAGGAGCTTATGAGCGAGCCCAT 1412  
Qy 1261 GAAGGTGTACAAAGACCGCGCAGGTCAATGAGTGTGAGAGTGAAGAGAGAGACCTT 1320  
Db 1413 GAAGGTGTACAAAGACCGCGCAGGTCAATGAGTGTGAGAGTGAAGAGAGAGACCTT 1472  
Qy 1331 CCGGAGAAAGTTTCAATGACGATCCCAAGAACTTTGGGCTGATCGATCTTCTGAGAG 1380  
Db 1473 CCGGAGAAAGTTTCAATGACGATCCCAAGAACTTTGGGCTGATCGATCTTCTGAGAG 1532  
Qy 1381 GAAGCAGTGTGAGTGTGCTCTCTATTACTAATGAGTGAAGAAATGAGAACTATTA 1440  
Db 1533 GAAGCAGTGTGAGTGTGCTCTCTATTACTAATGAGTGAAGAAATGAGAACTATTA 1592  
Qy 1441 GAGCTGTGAGACGAGAGCTATCGCGCGCGCGCAAGAGCGAGCAACACAGCAGCA 1500  
Db 1593 GAGCTGTGAGACGAGAGCTATCGCGCGCGCGCAAGAGCGAGCAACACAGCAGCA 1652  
Qy 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1560  
Db 1653 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1712  
Qy 1561 AGATGAGAAAGAGAGAAAGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
Db 1713 AGATGAGAAAGAGAGAAAGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1772  
Qy 1621 CGACAGAGAGAGACCTCTCTCAAGAGAGAGCAGACGACCTCAGGAGAGAGCAACGACGA 1680  
Db 1773 CGACAGAGAGAGACCTCTCTCAAGAGAGAGCAGACGACCTCAGGAGAGAGCAACGACGA 1832  
Qy 1681 GAAGAGGCTGTGGCTCTCAAGAGCGCAAACTGCGCAACGCGAGAGAGAGCGCAAGG 1740  
Db 1833 GAAGAGGCTGTGGCTCTCAAGAGCGCGCAAACTGCGCAACGCGAGAGAGAGCGCAAGG 1892  
Qy 1741 CGGATCAACCGCTCAATGGCTAATGAGGCGCAACAGCGAGAGAGGCGATCACCCCGAGA 1800  
Db 1893 CGGATCAACCGCTCAATGGCTAATGAGGCGCAACAGCGAGAGAGGCGATCACCCCGAGA 1952  
Qy 1801 GAGCGCCGAGCTGAGCTCTCATGAGCTGAATGAGAGTTCTCGCTGAGCAAGAAAGAAAT 1860  
Db 1953 GAGCGCCGAGCTGAGCTCTCATGAGCTGAATGAGAGTTCTCGCTGAGCAAGAAAGAAAT 2012  
Qy 1861 GGAACACGCGAAAGAGTCTCTGGAACAAGCGCGCAACGCTGCGGCAATGCGCCGAGT 1920  
Db 2013 GGAACACGCGAAAGAGTCTCTGGAACAAGCGCGCAACGCTGCGGCAATGCGCCGAGT 2072  
Qy 1921 GGTGGGCTTCAAGACTGTGTGCGAGGTAAAGAACTTCTACTTCAACTCAAGAAAGGCA 1980  
Db 2073 GGTGGGCTTCAAGACTGTGTGCGAGGTAAAGAACTTCTACTTCAACTCAAGAAAGGCA 2132  
Qy 1981 GAACCTCGATGAGATCTTGCAAGCAGACAAAGCTGAAGATGAGAAAGAGAGAAAGCGCG 2040  
Db 2133 GAACCTCGATGAGATCTTGCAAGCAGACAAAGCTGAAGATGAGAAAGAGAGAAAGCGCG 2192  
Qy 2041 GAGGAAGAAAGAAAGAGCGCGCGCGCGCGCGAGAGAGGCTGCATTCGCCGCCGTGGT 2100

Db 2193 GAGGAAGAAAGAAAGAGCGCGCGCGCGCGCGAGAGAGGCTGCATTCGCCGCCGTGGT 2252  
Qy 2101 GAGAGATGAGAGATGAGAGCGTCTGAGCGTGAAGCGAAATGAGAGAGATGATGAGAG 2160  
Db 2253 GAGAGATGAGAGATGAGAGCGTCTGAGCGTGAAGCGAAATGAGAGAGATGATGAGAG 2312  
Qy 2161 GAGTGAAGCTTTACATGCGCTCTGGGAATGAGGTGCCAGAGGGAGAAATGACGTGCCAGC 2220  
Db 2313 GAGTGA-----AGC 2321  
Qy 2221 CACTGTCAACAACAGCTGAGACACCGAGAGATCCCTCTCTCTCACTGAGAGCGCGCA 2280  
Db 2322 CACTGTCAACAACAGCTGAGACACCGAGAGATCCCTCTCTCTCACTGAGAGCGCGCA 2381  
Qy 2281 GGAACAAGGAGAGATGAGGCGCGCAAGCGCAAGCGCAAGCGCGCGCGCGCGCGCGCG 2340  
Db 2382 GGAACAAGGAGAGATGAGGCGCGCAAGCGCGCAAGCGCAAGCGCGCGCGCGCGCGCG 2441  
Qy 2341 AGGCCACCCACCCCAACCGAGAGACATCCCGGCGCGCGCGCGCGCGCGCGCGCGCTC 2400  
Db 2442 AGGCCACCCACCCCAACCGAGAGACATCCCGGCGCGCGCGCGCGCGCGCGCGCGCTC 2501  
Qy 2401 TGAAGCGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2460  
Db 2502 TGAAGCGACCTTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2561  
Qy 2461 TGTGTCTCCCAAGAGAGAGAGAGAGAGACCGCAGCAGCGCGCGCGCGCGCGCGCG 2520  
Db 2562 TGTGTCTCCCAAGAGAGAGAGAGAGAGACCGCAGCAGCGCGCGCGCGCGCGCGCG 2621  
Qy 2521 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
Db 2622 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2681  
Qy 2581 GCGCGTCAAGAGCGAGTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
Db 2682 GCGCGTCAAGAGCGAGTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2741  
Qy 2641 GAGAGCGCTGAGAGCGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
Db 2742 GAGAGCGCTGAGAGCGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2801  
Qy 2701 CGGAGAGCGCTCACTGCGCAAGAGCTCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 2760  
Db 2802 CGGAGAGCGCTCACTGCGCAAGAGCTCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 2861  
Qy 2761 CTGCAGTGCAAGCAGAGTGAATGAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
Db 2862 CTGCAGTGCAAGCAGAGTGAATGAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 2921  
Qy 2821 AAGGCCAGCTCTCACTGCGCAAGAGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 2880  
Db 2922 AAGGCCAGCTCTCACTGCGCAAGAGCTGCGAGCGCGCGCGCGCGCGCGCGCGCG 2981  
Qy 2881 ACTGAGCTTGAAGCAGTGAAGCAGCGAGCGCTGCGATTCCTCCCAATGAGGTCAACA 2940  
Db 2982 ACTGAGCTTGAAGCAGTGAAGCAGCGAGCGCTGCGATTCCTCCCAATGAGGTCAACA 3041  
Qy 2941 AGTCCATGAGCGCGCGCGAGAGAGCGAGCTCCCAAGAGAGAGAGAGAGAGAGAGAG 3000  
Db 3042 AGTCCATGAGCGCGCGCGAGAGAGCGAGCTCCCAAGAGAGAGAGAGAGAGAGAGAG 3101  
Qy 3001 GCACCGCGCAAACTGACGCGAGAGAGAGCGCGCTTCAAGAGCTTGAAGAGCGCGCG 3060  
Db 3102 GCACCGCGCAAACTGACGCGAGAGAGAGCGCGCTTCAAGAGCTTGAAGAGCGCGCG 3161  
Qy 3061 GGGCAAGAGCAGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3100  
Db 3162 GGGCAAGAGCAGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3221  
Qy 3101 ----GCGTTCAGCGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3156  
Db 3222 CCAGCTTTCGAGCGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3281

QY 3157 CTTGCCCTTCCCGTGGCCCCCGTGGAGTGAATCAAGGGCTTCCCGCATGCCCCGAGACC 3216  
Db 3282 CTGGCCCTTCCCGTGGCCCCCGTGGAGTGAATCAAGGGCTTCCCGCATGCCCCGAGACC 3341  
QY 3217 CTGAGCTTCTCTCAAGCTTCAACCTGTGTCAACCACTGGCCCTTGGGCTTCATGACATGCG 3276  
Db 3342 CTGAGCTTCTCTCAAGCTTCAACCTGTGTCAACCACTGGCCCTTGGGCTTCATGACATGCG 3401  
QY 3277 CCGGCCCCGCTTCCCGTGGCCCCCGTGGAGTGAATCAAGGGCTTCCCGCATGCCCCGAGACC 3336  
Db 3402 CCGGCCCCGCTTCCCGTGGCCCCCGTGGAGTGAATCAAGGGCTTCCCGCATGCCCCGAGACC 3461  
QY 3337 CAAGCAACCCGAGGCTCTCGAGAGGCAATATAGGTGCATCTCCCAAGAAATGTCCGCTCA 3396  
Db 3462 CAAGCAACCCGAGGCTCTCGAGAGGCAATATAGGTGCATCTCCCAAGAAATGTCCGCTCA 3521  
QY 3397 GCTCAAGTCCGTAATCAAGAGATGCAAGAGCCCGGTGGGCTTGTCAACAATGGGCT 3456  
Db 3522 GCTCAAGTCCGTAATCAAGAGATGCAAGAGCCCGGTGGGCTTGTCAACAATGGGCT 3581  
QY 3457 GCGCCCTGCTTGAAGACCCCAAAAAGCTGGCACTTTCAGCGAGTGAAGCAGAGAGCT 3516  
Db 3582 GCGCCCTGCTTGAAGACCCCAAAAAGCTGGCACTTTCAGCGAGTGAAGCAGAGAGCT 3641  
QY 3517 GTCCCAACGGGGGCGAGGCTGGGCAACCGAGAGCTGGGGGTGCCAAGCCGAGAGGCG 3576  
Db 3642 GTCCCAACGGGGGCGAGGCTGGGCAACCGAGAGCTGGGGGTGCCAAGCCGAGAGGCG 3701  
QY 3577 GTCCGTGTGAAGAGGACGCTCTGGGCTCAAGTTCGGGGCGAGAGATCAACAAAGGCAAT 3636  
Db 3702 GTCCGTGTGAAGAGGACGCTCTGGGCTCAAGTTCGGGGCGAGAGATCAACAAAGGCAAT 3761  
QY 3637 TCCCAAGCAACGGGGTCCCTGCGACAGCGCATCAATACCGGGCTCCATCAACCAAGG 3696  
Db 3762 TCCCAAGCAACGGGGTCCCTGCGACAGCGCATCAATACCGGGCTCCATCAACCAAGG 3821  
QY 3697 CAGCGAGCTGAAGTCTGTGAAGGGGCAACATCAAGAGTCAATGGGGGAGAGACGCC 3756  
Db 3822 CAGCGAGCTGAAGTCTGTGAAGGGGCAACATCAAGAGTCAATGGGGGAGAGACGCC 3881  
QY 3757 GAGTCTTGTGAACGGGGCGGAGAGACAGCTTCCCAAGGGCAAGTCAATCAAGAGG 3816  
Db 3882 GAGTCTTGTGAACGGGGCGGAGAGACAGCTTCCCAAGGGCAAGTCAATCAAGAGG 3941  
QY 3817 CAAGAAAGGCGACGCTTGTCTTAAGAGGTGCATGTCTGTGAACCAAGTCTCAAGAG 3876  
Db 3942 CAAGAAAGGCGACGCTTGTCTTAAGAGGTGCATGTCTGTGAACCAAGTCTCAAGAG 4001  
QY 3877 GAGCGGCAAGAGAGTCAAGAGACCCGCCATGAAGGGCGCGCCCAAGGGCACTTAAG 3936  
Db 4002 GAGCGGCAAGAGAGTCAAGAGACCCGCCATGAAGGGCGCGCCCAAGGGCACTTAAG 4061  
QY 3937 CATGATGAGAGGCGGCTGGGAGAGAGCAATCTCTCAAGCAGCATGAAAGTCTCATGGG 3996  
Db 4062 CATGATGAGAGGCGGCTGGGAGAGAGCAATCTCTCAAGCAGCATGAAAGTCTCATGGG 4121  
QY 3997 CCGTGCATATCCGCGCGAGAGCAACAGCCCAACCTTCAAGAGAGCAACCATCCG 4056  
Db 4122 CCGTGCATATCCGCGCGAGAGCAACAGCCCAACCTTCAAGAGAGCAACCATCCG 4181  
QY 4057 CCGGTTCATCAACAAGAGATCCCTGGTCTTAAGAGAGGCAAGAGAGCATACCTGGG 4116  
Db 4182 CCGGTTCATCAACAAGAGATCCCTGGTCTTAAGAGAGGCAAGAGAGCATACCTGGG 4241  
QY 4117 TCGGAGGCGCAAGCTCTAAAGCGGAGAGGAGCGCTTCCGCCCCCAAGCGGCTCAAGG 4176  
Db 4242 TCGGAGGCGCAAGCTCTAAAGCGGAGAGGAGCGCTTCCGCCCCCAAGCGGCTCAAGG 4301  
QY 4177 CTTGACCGAGGCTTAACAAGCGAGGCTTGGGCTTGAAGCTGAAGCGGCGCATGA 4236  
Db 4302 CTTGACCGAGGCTTAACAAGCGAGGCTTGGGCTTGAAGCTGAAGCGGCGCATGA 4361

QY 4237 GGGCTGTGTGGCCACGGTGAAGAGGCGGGCGCTCATCTCATAGAGATCCCGCGAGAG 4296  
Db 4362 GGGCTGTGTGGCCACGGTGAAGAGGCGGGCGCTCATCTCATAGAGATCCCGCGAGAG 4421  
QY 4297 GCTGCGGCAACCGCCGAGGTGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGG 4356  
Db 4422 GCTGCGGCAACCGCCGAGGTGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGG 4481  
QY 4357 GCAGGGGCAACCGGCTCAAGTGAAGCAACGGCGGCTCAACCACTGGGCTCAAAAAGCAG 4416  
Db 4482 GCAGGGGCAACCGGCTCAAGTGAAGCAACGGCGGCTCAACCACTGGGCTCAAAAAGCAG 4541  
QY 4417 CGTACGCTCTCTCATCGGAGAGCCCGCGGAGAGCTTCCACCCGTGACCCCGCTGATGT 4476  
Db 4542 CGTACGCTCTCTCATCGGAGAGCCCGCGGAGAGCTTCCACCCGTGACCCCGCTGATGT 4601  
QY 4477 GATGCGGAGCGCCGGGCACTGGAAGTGTCTGCTAAGAGAGGCTTGAAGCGGCGC 4536  
Db 4602 GATGCGGAGCGCCGGGCACTGGAAGTGTCTGCTAAGAGAGGCTTGAAGCGGCGC 4661  
QY 4537 AGGGAACGGCAAGAGCTCGGGGGGCTCAATGGCGGGGGCGCGGCTCATTTGGCTGA 4596  
Db 4662 AGGGAACGGCAAGAGCTCGGGGGGCTCAATGGCGGGGGCGCGGCTCATTTGGCTGA 4721  
QY 4597 GCTGGGTAAGCGCGGAGAGAGCCCGCTGAACCTATGAGAGCAACAGGAGCACTTTGCGG 4656  
Db 4722 GCTGGGTAAGCGCGGAGAGAGCCCGCTGAACCTATGAGAGCAACAGGAGCACTTTGCGG 4781  
QY 4657 CCACTTCCACAGAGGTTCCGCGGTGAACATGGGGAGCCCAACCGCGCTGCAAGAGG 4716  
Db 4782 CCACTTCCACAGAGGTTCCGCGGTGAACATGGGGAGCCCAACCGCGCTGCAAGAGG 4841  
QY 4717 CAGCTTTTGTCTCAAGAGAGCATCCAGAGCCGAAGCTGAAGCTCAAGCTGTGAGAT 4776  
Db 4842 CAGCTTTTGTCTCAAGAGAGCATCCAGAGCCGAAGCTGAAGCTGTGAGAT 4901  
QY 4777 CGCAAGTCCCGGACAGAGACGTGCCGAGCAACCAACCCATCTCGCCTATGA 4836  
Db 4902 CGCAAGTCCCGGACAGAGACGTGCCGAGCAACCAACCCATCTCGCCTATGA 4961  
QY 4837 GCACTGTCTTGGGGGCTGAAGTGGCTGTGAACCTGTATCGAGCAACATCCCTGGCTT 4896  
Db 4962 GCACTGTCTTGGGGGCTGAAGTGGCTGTGAACCTGTATCGAGCAACATCCCTGGCTT 5021  
QY 4897 CCAACCCCACTCAATPCCCCGGGGCTTCCCTCTGAGACGAGCGCTGCTTAATCACTGGC 4956  
Db 5022 CCAACCCCACTCAATPCCCCGGGGCTTCCCTCTGAGACGAGCGCTGCTTAATCACTGGC 5081  
QY 4957 CCGACACTGGGCCCGCAACCCCACTTACCGGACCTGTACCCACCTCATCCGCGG 5016  
Db 5082 CCGACACTGGGCCCGCAACCCCACTTACCGGACCTGTACCCACCTCATCCGCGG 5141  
QY 5017 CTACCCCGACAGCGCGGCTGAGAGACCGGAGACATCATCAATGACTCATCACTTC 5076  
Db 5142 CTACCCCGACAGCGCGGCTGAGAGACCGGAGACATCATCAATGACTCATCACTTC 5201  
QY 5077 GCAGCAGATGACCAACAACCGGCAACCGGCAATGGGCTGAGAGCTGATGTGAGGGG 5136  
Db 5202 GCAGCAGATGACCAACAACCGGCAACCGGCAATGGGCTGAGAGCTGATGTGAGGGG 5261  
QY 5137 CCTCTGCGCCGAGAGTCTCTGTGCACTCAATGAGTGGCGGGTCCCGAGGACATCAT 5196  
Db 5262 CCTCTGCGCCGAGAGTCTCTGTGCACTCAATGAGTGGCGGGTCCCGAGGACATCAT 5321  
QY 5197 CGACTGTCCAAAGTGCACACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 5256  
Db 5322 CGACTGTCCAAAGTGCACACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 5381  
QY 5257 CACCGGCAATGAGACCGGCTTGGCTCACTCCCAACCGGCGCCGAGCCCTTGAAGAGCGGCA 5316  
Db 5382 CACCGGCAATGAGACCGGCTTGGCTCACTCCCAACCGGCGCCGAGCCCTTGAAGAGCGGCA 5441  
QY 5317 CAGCACTCCCACTTCCCGAGAGGTCAACCACTTGAACAAAACCAACCAACCACTGTC 5376



Db	5442	CAGAGACGTCCCCACTCTCTCCAGGAGGTCCAACACTTTGACAAACCAACCAACGCTC	5501
Qy	5377	CTCGTCCGAGCGGGAGCGGAGCCGGAGATCGAAGCGGGAGCCGGATTCGGAGCGGAAAA	5436
Db	5502	CTCGTCCGAGCGGGAGCGAGACCGGGATCGAAGCGGGAGCCGGAGTCGGGAGCGGAAAA	5561
Qy	5437	GTCATCTCTCAACGTTCACACAGACGGTGTGACACGACCTCATCTGTGACACTGTGTACGA	5496
Db	5562	GTCATCTCTCAACGTTCACACAGACGGTGTGACACGACCTCATCTGTGACACTGTGTACGA	5621
Qy	5497	GCAGAGCAGCGGACGAGCGGACGAGCGGGGGGTGGGGGACAGCAGCCGCCCGC	5556
Db	5522	GCAGAGCAGCGGACGAGCGGACGAGCGGGGGGTGGGGGACAGCAGCCGCCCGC	5681
Qy	5557	CTCCCACTCCCATGSCCCACAGACTCGGCCATCTCCCTCGAACCCAGATGCGCTTCA	5616
Db	5582	CTCCCACTCCCATGSCCCACAGACTCGGCCATCTCCCTCGAACCCAGATGCGCTTCA	5741
Qy	5617	GCAGAGACCCAGTGTGCTTCAACACAGGACATGAGGGTATCATCACCGCTGTGAGACC	5676
Db	5742	GCAGAGACCCAGTGTGCTTCAACACAGGACATGAGGGTATCATCACCGCTGTGAGACC	5801
Qy	5677	CAGCAAGCGCCAGGTCCTGAGGTCAACCTCAACCTCCACCGGTGCGCCAGCGCCAC	5736
Db	5802	CAGCAAGCGCCAGGTCCTGAGGTCAACCTCAACCTCCACCGGTGCGCCAGCGCCAC	5861
Qy	5737	ATTCCCACTTGACACCCACCTGCCACTGGGCGGCACTCTGATGGGGTCTACCTTACCT	5796
Db	5862	ATTCCCACTTGACACCCACCTGCCACTGGGCGGCACTCTGATGGGGTCTACCTTACCT	5921
Qy	5797	CATGGAAGCCGCTCTTGCTGCTCCAAAGAGAGCCCCCGGGTGCCTCGGACAGAGCGGCCG	5856
Db	5922	CATGGAAGCCGCTCTTGCTGCTCCAAAGAGAGCCCCCGGGTGCCTCGGACAGAGCGGCCG	5981
Qy	5857	AGCAGACACCGGGCATGCTTCTCTCGCCAAACCCCGACCGGCTCCGGGCTGAGAGCCCGC	5916
Db	5982	AGCAGACACCGGGCATGCTTCTCTCGCCAAACCCCGACCGGCTCCGGGCTGAGAGCCCGC	6041
Qy	5917	CTCTCTCCCGCAGCAAGGGCTCGAGAGCCCGGACCCTAGTGCCTCTGTCTCTGAGCAGC	5976
Db	6042	CTCTCTCCCGCAGCAAGGGCTCGAGAGCCCGGACCCTAGTGCCTCTGTCTCTGAGCAGC	6101
Qy	5977	CACCATGCGCCGCAACCCCTGCGAAAGACTTCGACACTTCACAGCGCAGACCCGAGCCGC	6036
Db	6102	CACCATGCGCCGCAACCCCTGCGAAAGACTTCGACACTTCACAGCGCAGACCCGAGCCGC	6161
Qy	6037	GGCGCCACCTGECCTCGGCTCGGACCCGCAACCGGAAAAAGCTCAAGTAAACCTTTTC	6096
Db	6162	GGCGCCACCTGECCTCGGCTCGGACCCGCAACCGGAAAAAGCTCAAGTAAACCTTTTC	6221
Qy	6097	CATCAGAGAACTGGAATCCCGTTCTCTGAGGTAAACAAGGAGCAGACTACAGCCCGAAG	6156
Db	6222	CATCAGAGAACTGGAATCCCGTTCTCTGAGGTAAACAAGGAGCAGACTACAGCCCGAAG	6281
Qy	6157	GGTGAAGCCCGTCAAGCCCTGTGAGCTCAACCAAGTCTGACCCACGACAAAGGGCTCCCA	6216
Db	6282	GGTGAAGCCCGTCAAGCCCTGTGAGCTCAACCAAGTCTGACCCACGACAAAGGGCTCCCA	6341
Qy	6217	GCACCTGGAAGACTTCGACAAAGCACACTTGAAGGGGAGCTGGGGCCCAAGCAGCCAG	6276
Db	6342	GCACCTGGAAGACTTCGACAAAGCACACTTGAAGGGGAGCTGGGGCCCAAGCAGCCAG	6401
Qy	6277	CCCGGTGAAGCTTGGCGGGGAGGCGCCCACTCCACACCTGCGGCGCGCTGAGAG	6336
Db	6402	CCCGGTGAAGCTTGGCGGGGAGGCGCCCACTCCACACCTGCGGCGCGCTGAGAG	6461
Qy	6337	CCAGCCCTCTGTCAGACCCTGCTTCAGACCGCCCAAGGGGTCAAAAGTTCACAGCGGGT	6396
Db	6462	CCAGCCCTCTGTCAGACCCTGCTTCAGACCGCCCAAGGGGTCAAAAGTTCACAGCGGGT	6521
Qy	6397	GGTCAACCTGGCCCAAGCACAATAGTGAAGTCATCAACAGACTTACACCCGACACACC	6456

Dp	6522	GGTCACTCCCTGGCCCGACATCAAGTAGAGTTCATCACAACAAGACTCAACACCGGGACACACCC	6581
Qy	6457	ACAGCAAGCTGAGGGGCATCCCTGTCGGCCGCCCTCTTACTCTTCTCTGGGGGCAGCTGGCC	6516
Dp	6582	ACAGCAAGCTGAGGGGCATCCCTGTCGGCCGCCCTCTTACTCTTCTCTGGGGGCAGCTGGCC	6641
Qy	6517	CGTCTGGAAGCTCCGGCGGGCCCAACCAATGACCTCTACTCCCGGCCCGGACCATGTGAC	6576
Dp	6642	CGTCTGGAAGCTCCGGCGGGCCCAACCAATGACCTCTACTCCCGGCCCGGACCATGTGAC	6701
Qy	6577	CCCGGGCCCGTGGCTCCCCCGACAGCGAAGGGGGGCAAGGCTTTCAGAGCCAAACAAGAC	6636
Dp	6702	CCCGGGCCCGTGGCTCCCCCGACAGCGAAGGGGGGCAAGGCTTTCAGAGCCAAACAAGAC	6761
Qy	6637	GTCGGTCTTGGGTGGTGGTAGAGCGGTATTGAACTGTGTCCCCACCGAGGGCATATAC	6696
Dp	6762	GTCGGTCTTGGGTGGTGGTAGAGCGGTATTGAACTGTGTCCCCACCGAGGGCATATAC	6821
Qy	6697	GGAGCCAGGGGCATCTCCGGAGTGTGTGTACCCGCTGTGTACCCGGGATGGGGAAACGAC	6756
Dp	6822	GGAGCCAGGGGCATCTCCGGAGTGTGTGTACCCGCTGTGTACCCGGGATGGGGAAACGAC	6881
Qy	6757	GGAGCCCGAGGAGATGGGGCTCCAAAGTCTCCAGGGCAACACAGCCAGCGCCAGCTTCTT	6816
Dp	6882	GGAGCCCGAGGAGATGGGGCTCCAAAGTCTCCAGGGCAACACAGCCAGCGCCAGCTTCTT	6941
Qy	6817	CAGCAAGCTGACCGGAGAGCACTCCGCGATGTCTCAAGTCCAGAGCAAGAGCAAGATCAACAA	6876
Dp	6942	CAGCAAGCTGACCGGAGAGCACTCCGCGATGTCTCAAGTCCAGAGCAAGAGATCAACAA	7001
Qy	6877	GAACTGTAAACCCCAACAACCGGAATGAGCCTGAATATCAATATCAGCCACGCTGGGAACGA	6936
Dp	7002	GAACTGTAAACCCCAACAACCGGAATGAGCCTGAATATCAATATCAGCCACGCTGGGAACGA	7061
Qy	6937	GATTTCAATATGCGCGCGCATCAACCGGAACGCGCTTATGACCTATGAAGCCAGCGCGT	6996
Dp	7062	GATTTCAATATGCGCGCGCATCAACCGGAACGCGCTTATGACCTATGAAGCCAGCGCGT	7121
Qy	6997	GCAAGGAACATGCGCAGCAACAATGGGGCTTGAAGGCCATATATGAAAGGCATCATGGG	7056
Dp	7122	GCAAGGAACATGCGCAGCAACAATGGGGCTTGAAGGCCATATATGAAAGGCATCATGGG	7181
Qy	7057	TAAATATGACCATGGGGAAGATCCCCGCGCTCAGGGCCAAATGCTTTTAACTCTTGAA	7116
Dp	7182	TAAATATGACCATGGGGAAGATCCCCGCGCTCAGGGCCAAATGCTTTTAACTCTTGAA	7241
Qy	7117	TGCGAGTGCAGCGCTGCGCTGTATAGCCCATTAACCGCTGTGACGGAAGGATGACCA	7176
Dp	7242	TGCGAGTGCAGCGCTGCGCTGTATAGCCCATTAACCGCTGTGACGGAAGGATGACCA	7301
Qy	7177	CACACTCACTTCGCGCAGATGGCGCGGAGGAAGGCACAAAGTCTTGGCAGACCCACAGCACCG	7236
Dp	7302	CACACTCACTTCGCGCAGATGGCGCGGAGGAAGGCACAAAGTCTTGGCAGACCCACAGCACCG	7361
Qy	7237	AAAAGCCAAGTCCCCCGGCCCCCGGGGCTTGGCATTTGGGGAACCGGCACCTCTGTCTCTC	7296
Dp	7362	AAAAGCCAAGTCCCCCGGCCCCCGGGGCTTGGCATTTGGGGAACCGGCACCTCTGTCTCTC	7421
Qy	7297	AGTGCATCTCGGAGGGAGAACTGCAACCGCGCGGAGCGCGCTCAACCAACCGCGTGGAGGGA	7356
Dp	7422	AGTGCATCTCGGAGGGAGAACTGCAACCGCGCGGAGCGCGCTCAACCAACCGCGTGGAGGGA	7481
Qy	7357	CAGGCACTTCGTCGCGCAGATTTCAAGCAATTCCTCTTAAACCCCTGTATCAATGGCGCTGCA	7416
Dp	7482	CAGGCACTTCGTCGCGCAGATTTCAAGCAATTCCTCTTAAACCCCTGTATCAATGGCGCTGCA	7541
Qy	7417	GGCGGGGTGATGGCTTCCCAACCCCAACCGGAGCTCCCGCGGGACAGCGGCCCCCTTGC	7476
Dp	7542	GGCGGGGTGATGGCTTCCCAACCCCAACCGGAGCTCCCGCGGGACAGCGGCCCCCTTGC	7601
Qy	7477	TGGCCCCCAACAAGCTCTGGGACGAGGAGCCCAAGCACTGTCTGTCTGCAATACGAGAC	7536
Dp	7602	TGGCCCCCAACAAGCTCTGGGACGAGGAGCCCAAGCACTGTCTGTCTGCAATACGAGAC	7661

```

QY 7537 ACTCTCCGACAGGAGTGACTCAGAAACAGGGCGGGGGGGCGGGCGGTUERSPERLUG 7596
      |||||
      7662 ACTCTCCGACAGGAGTGACTCAGAAACAGGGCGGGGGGGGGCGGGCGG-----TG 7711
QY 7597 TCAGGTCCGACGAGCCACAGAAACGGCCCTTGACAGAGCGGGCGGCTGCCATCTCCCC 7656
      |||||
      7712 TCAGGTCCGACGAGCCACAGAAACGGCCCTTGACAGAGCGGGCGGCTGCCATCTCCCC 7771
QY 7657 AACCAAGAGAGAGAGCCCTGAGTCCCGCTGCGCTCCATCTGTCCTCCAGAGCCG 7716
      |||||
      7772 AACCAAGAGAGAGAGCCCTGAGTCCCGCTGCGCTCCATCTGTCCTCCAGAGCCG 7831
QY 7717 GCATCCCTGCTCTCTAAAGCCTTAAAGCTTAAAGCTCCCGCCCGGGCGGCTGTGAGA 7776
      |||||
      7832 GCATCCCTGCTCTCTAAAGCCTTAAAGCTTAAAGCTCCCGCCCGGGCGGCTGTGAGA 7891
QY 7777 CCTTACTCAGGGAGATGTTTACTGTGCTCGGAGAGGAGGAGGAGGCGGGAGAGGG 7836
      |||||
      7892 CCTTACTCAGGGAGATGTTTACTGTGCTCGGAGAGGAGGAGGAGGCGGGAGAGGG 7951
QY 7837 GCACGGCAGGCGGTGTGAGCAGCCACACAGGCGGGCGGCGAGGAGCCAAAGCAG 7896
      |||||
      7952 GCACGGCAGGCGGTGTGAGCAGCCACACAGGCGGGCGGCGAGGAGCCAAAGCAG 8011
QY 7897 GATGACACGACCTCAGGCGCAGCTGCTCCCGGATGCAATTGGAACCAAGCTTAA 7956
      |||||
      8012 GATGACACGACCTCAGGCGCAGCTGCTCCCGGATGCAATTGGAACCAAGCTTAA 8071
QY 7957 CTGAGCTCGCAGGCGCCCGCGCCCTCCCTCCGCTCCCATCCCTTAGCGCTCGACAG 8016
      |||||
      8072 CTGAGCTCGCAGGCGCCCGCGCCCTCCCTCCGCTCCCATCCCTTAGCGCTCGACAG 8131
QY 8017 ATGAGACGACGCGCTCTGACGCCCGCAGTCCGCTCGTCCGGTCCCAACAGCTGCCCA 8076
      |||||
      8132 ATGAGACGACGCGCTCTGACGCCCGCAGTCCGCTCGTCCGGTCCCAACAGCTGCCCA 8191
QY 8077 GCCAAGGAGTTCGTAAGAACCAAGTCAAGGCGAGGCGGAGGAGCAAAAGGCGCAGGTCG 8136
      |||||
      8192 GCCAAGGAGTTCGTAAGAACCAAGTCAAGGCGAGGCGGAGGAGCAAAAGGCGCAGGTCG 8251
QY 8137 GCCTGGGGGAGAGCGAGTGTCCGAGAGCTGAGCTGTTTTTTCACATCGTTCGCGCAG 8196
      |||||
      8252 GCCTGGGGGAGAGCGAGTGTCCGAGAGCTGAGCTGTTTTTTCACATCGTTCGCGCAG 8311
QY 8197 CGGTGGGAGAGAGCGAGTGTAAATGATGTGTGTTTTCACAGGATATTTTGTATAC 8256
      |||||
      8312 CGGTGGGAGAGAGCGAGTGTAAATGATGTGTGTTTTCACAGGATATTTTGTATAC 8371
QY 8257 CTTCATGAAATTAATTCAGATGTTTTTACGCAAGAGAGGACTTACCAATTAATTCGTGC 8316
      |||||
      8372 CTTCATGAAATTAATTCAGATGTTTTTACGCAAGAGAGGACTTACCAATTAATTCGTGC 8431
QY 8317 TGTGCTTTTGTATCTGCTTACCGTTCAAGAGGCGGTGAGGCGCAGACGTCCGTGACC 8376
      |||||
      8432 TGTGCTTTTGTATCTGCTTACCGTTCAAGAGGCGGTGAGGCGCAGACGTCCGTGACC 8491
QY 8377 CATCACTCCGAGAGCAAGGAGGCGGAGACTGCTGTCAAGCGCCGCTGTCTCTCCCTC 8436
      |||||
      8492 CATCACTCCGAGAGCAAGGAGGCGGAGACTGCTGTCAAGCGCCGCTGTCTCTCCCTC 8551
QY 8437 CCGCCCTCTCTGGGAGAGATGAAATTCAGATGCTATTCGTGCGCCGCAATTCGCGAGGG 8496
      |||||
      8552 CCGCCCTCTCTGGGAGAGATGAAATTCAGATGCTATTCGTGCGCCGCAATTCGCGAGGG 8611
QY 8497 TGTGCTTTTGTATCTGCTTACCAAGAGGCGGTGAGGCGCAGACGTCCGTGACC 8556
      |||||
      8612 TGTGCTTTTGTATCTGCTTACCAAGAGGCGGTGAGGCGCAGACGTCCGTGACC 8671
QY 8557 AAAAAAAAAAAAAA 8571
      |||||
      8672 AAAAAAAAAAAAAA 8686
      |||||

```

```

RESULT 11
ADL12577
ID ADL12577 standard; cDNA; 9053 BP.
XX
AC ADL12577;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human steroid-induced C3A liver cell cDNA #306.
XX
KW se; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;
XX
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
XX
OS Homo sapiens.
XX
PN US673549-B1.
XX
PD 06-JAN-2004.
XX
PF 12-OCT-2001; 2001US-00976594.
XX
PR 12-OCT-2000; 2000US-0240409P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Furness LM, Buchbinder JL;
XX
DR WPI; 2004-068610/07.
XX
PT Combination useful for preparing a composition for treating liver
PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT comprises cDNAs that are differentially expressed in response to steroid
PT treatment.
XX
PS Claim 1; SEQ ID NO 306; 141pp; English.
XX
XX
CC The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a standard defines a stage of that
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and at
CC least one molecule or compound. The molecules or compounds are regulatory
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human cDNA which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 9053 BP; 1904 A; 3171 C; 2722 G; 1254 T; 0 U; 2 Other;
Query Match 96.1%; Score 8226.8; DB 12; Length 9053;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 8451; Conservative 2; Mismatches 42; Indels 103; Gaps 10;
QY 1 CATGTGGGCTCCACACAGCTTGTGGCAACAGCTGAGGCGCAGTCCGCTGACCC 60
      |||||
      519 CATGTGGGATCCACACAGCTGTGGCAACAGCTGAGGCGCAGTCCGCTGACCC 578
QY 61 GCCCCACAGCTTTCTTACCCAGTGCAGATCGCCGAGAGCAACAGGAGTCCGAGCTCT 120
      |||||
      579 GCCCCACAGCTTTCTTACCCAGTGCAGATCGCCGAGAGCAACAGGAGTCCGAGCTCT 638
QY 121 GAGATACGACCACTCCCGGAGCTATGCTCCCACTGTGCGCGGAGCTCATATCCA 180
      |||||

```

Db 639 GAGATACCAACAACAACCTCCGCGACTATGCTCCACCTGTGCGCCGGCTCATATCCA 698  
Qy 181 GCCCAGCGCGCGAGAGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGAAATGAACGGTCCA 240  
Db 699 GCCCAGCGCGCGAGAGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGAAATGAACGGTCCA 758  
Qy 241 GAGGCTCCAACCTGCGGCGCAAGATCCCACTCATACCTGCGGAGCTGGGGAATCAGAGAT 300  
Db 759 GAGGCTCCAACCTGCGGCGCAAGATCCCACTCATACCTGCGGAGCTGGGGAATCAGAGAT 818  
Qy 301 GAGGTTCAATTGAAAGCAAGCGCGCTCGGCTAGAGCTGCGCTGACCCCGTCTGCGACC 360  
Db 819 GAGGTTCAATTGAAAGCAAGCGCGCTCGGCTAGAGCTGCGCTGACCCCGTCTGCGACC 878  
Qy 361 GTCAACCCCTGTGTGCGCAACGCGGCGCAAGCTTGAGAGACTTCAACCAAGAACCGTAA 420  
Db 879 GTCAACCCCTGTGTGCGCAACGCGGCGCAAGCTTGAGAGACTTCAACCAAGAACCGTAA 938  
Qy 421 CTTGACGGGCAAGCTTGGAAACGGGTGTCTTCCCCCAGCCCCCGCACTGACCTTGAGCT 480  
Db 939 CTTGACGGGCAAGCTTGGAAACGGGTGTCTTCCCCCAGCCCCCGCACTGACCTTGAGCT 998  
Qy 481 GGAAGCTGGTGCGGCAACGCGCTGTGCCAAGGAGGAGCTGATCCAGAACATTTGAAACGGGTGGA 540  
Db 999 GGAAGCTGGTGCGGCAACGCGCTGTGCCAAGGAGGAGCTGATCCAGAACATTTGAAACGGGTGGA 1058  
Qy 541 CCGAGAGATCAACCATGTGTAGACAGACAGATCTCTAAGCTTGAAGAAAGAACAGCAACAGCT 600  
Db 1059 CCGAGAGATCAACCATGTGTAGACAGACAGATCTCTAAGCTTGAAGAAAGAACAGCAACAGCT 1118  
Qy 601 GAGAGAGAGAGGCTGCCAAGCCGCGCGAGCTTGAGAGGCCGTGTGTCACCGCGCCATCGA 660  
Db 1119 GAGAGAGAGAGGCTGCCAAGCCGCGCGAGCTTGAGAGGCCGTGTGTCACCGCGCCATCGA 1178  
Qy 661 GTCCGAAGCACCGGAGCTGTGTGACATCATACGACGAGAAACCGGAAAGAGCTGAAGC 720  
Db 1179 GTCCGAAGCACCGGAGCTGTGTGACATCATACGACGAGAAACCGGAAAGAGCTGAAGC 1238  
Qy 721 TGCAATCGGATTTTGGAAGGCTGTGGGCGCCAGAGGTGAGCTGCGCTGTACACAGACC 780  
Db 1239 TGCAATCGGATTTTGGAAGGCTGTGGGCGCCAGAGGTGAGCTGCGCTGTACACAGACC 1298  
Qy 781 CTCGCAACACCGCGCGAGTATCATGAGAACATCAAAATATAACGAGCGATGCGGAGAGAGCT 840  
Db 1299 CTCGCAACACCGCGCGAGTATCATGAGAACATCAAAATATAACGAGCGATGCGGAGAGAGCT 1358  
Qy 841 AATCTTGTACTTCAAGAGGAGGATCAAGCTGCGGAAACAAATGGAAGCGAAGTTCTGCGCA 900  
Db 1359 AATCTTGTACTTCAAGAGGAGGATCAAGCTGCGGAAACAAATGGAAGCGAAGTTCTGCGCA 1418  
Qy 901 GCGCTATGACCAAGCTCATGAGAGGCTTTGGAAAAAAAGTGAGCGCATCGAAAAACAACC 960  
Db 1419 GCGCTATGACCAAGCTCATGAGAGGCTTTGGAAAAAAAGTGAGCGCATCGAAAAACAACC 1478  
Qy 961 GCGCGCGCGCGCGCAAGAGAGCAAGGTGCGGAGTACTACGAAAAACAGTTCCCTGAGAT 1020  
Db 1479 CCGGCGCGCGCGCGCAAGAGAGCAAGGTGCGGAGTACTACGAAAAACAGTTCCCTGAGAT 1538  
Qy 1021 CCGGAAGCAGCGCGAGCTGCAAGAGCGCATGCAAGAGGCGGCGCAAGCGGCGGAGAGTGG 1080  
Db 1539 CCGGAAGCAGCGCGAGCTGCAAGAGCGCATGCAAGAGGCGGCGCAAGCGGCGGAGAGTGG 1598  
Qy 1081 GCTGTTCATGTGTGCGCGCGCGCAAGAGCGAGATGATCAGAGATCATGATGAGCTCTGC 1140  
Db 1599 GCTGTTCATGTGTGCGCGCGCGCGCAAGAGCGAGATGATCAGAGATCATGATGAGCTCTGC 1658  
Qy 1141 AGAGCAGAGAAACTTGAAGAGACAGATGCGCAAGCTGCGCTGTATCCCGCATGCTGTA 1200  
Db 1659 AGAGCAGAGAAACTTGAAGAGAGACAGATGCGCAAGCTGCGCTGTATCCCGCATGCTGTA 1718  
Qy 1201 CGAGCCTCAACCAAGCGCGCATGAGTTTCAATCAATGAACGGGCTTTATGGCCGACCCAT 1260  
Db 1719 CGAGCCTCAACCAAGCGCGCATGAGTTTCAATCAATGAACGGGCTTTATGGCCGACCCAT 1778

Qy 1261 GAAAGTGTACAAAGACCGCCAGGTATGAACATGTGAGTGAAGCAGAGAAAGAACCTT 1320  
Db 1779 GAAAGTGTACAAAGACCGCCAGGTATGAACATGTGAGTGAAGCAGAGAAAGAACCTT 1838  
Qy 1321 CCGGAGAAAGTTTCAATGACGATCCCAAGAACTTTGGCCTGTATCGATCATTTCTTGAAG 1380  
Db 1839 CCGGAGAAAGTTTCAATGACGATCCCAAGAACTTTGGCCTGTATCGATCATTTCTTGAAG 1898  
Qy 1381 GAAAGCAGTGGCTGAGTGGCTCTCTTACTTACTTACTTAAAGAAATGAGAACTTATA 1440  
Db 1899 GAAAGCAGTGGCTGAGTGGCTCTCTTACTTACTTACTTAAAGAAATGAGAACTTATA 1958  
Qy 1441 GAGCCTGTGAGACGAGACTATCGGCGCGCGCAAGAGCGACAGCAACA--CAGCA 1497  
Db 1959 GAGCCTGTGAGACGAGACTATCGGCGCGCGCGCAAGAGCGAGCAACACAGCA 2018  
Qy 1498 GCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1557  
Db 2019 GCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2078  
Qy 1558 GAAAGATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1617  
Db 2079 GAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2138  
Qy 1618 GAAACGACAAAGAGACCTCTTCAAGAGAGAGACAGACGACCTTCAAGGAGAGACAA 1677  
Db 2139 GAAACGACAAAGAGACCTCTTCAAGAGAGAGACAGACGACCTTCAAGGAGAGACAA 2198  
Qy 1678 CGAGAGAGAGGCTGTGGCTCCAAAGGCGCGCAAACTGCCCAACGCGCGAGAGACGCA 1737  
Db 2199 CGAGAGAGAGGCTGTGGCTCCAAAGGCGCGCAAACTGCCCAACGCGCGAGAGACGCA 2258  
Qy 1738 AGGCGGCATCAACCGCTCAATGAGCTTAATGAGGCGCAACGAGAGAGGCGCATCACCCCA 1797  
Db 2259 AGGCGGCATCAACCGCTCAATGAGCTTAATGAGGCGCAACGAGAGAGGCGCATCACCCCA 2318  
Qy 1798 GCAGAGCGCGGAGCTGGCTTCATGAGAGCTGATGAGAGTTCTGTGCGAGACAAAGAA 1857  
Db 2319 GCAGAGCGCGGAGCTGGCTTCATGAGAGCTGATGAGAGTTCTGTGCGAGACAAAGAA 2378  
Qy 1858 AATGGAACACGCGCAAGAAAGGTTCTCTTGAACACGCGCGCAACTGTGTGGCCATGCGCCG 1917  
Db 2379 AATGGAACACGCGCAAGAAAGGTTCTCTTGAACACGCGCGCAACTGTGTGGCCATGCGCCG 2438  
Qy 1918 GATGCTGGGCTCCAAAGCTGTGTGCGAGTGTAAAGACTTCTTCACTTCACTCAAGAAAG 1977  
Db 2439 GATGCTGGGCTCCAAAGCTGTGTGCGAGTGTAAAGACTTCTTCACTTCACTCAAGAAAG 2498  
Qy 1978 GCAGAACCTGATGAGATCTTGCAGCAGCACAGCTGAAATGAGAGAGAGAGAGAGAGAG 2037  
Db 2499 GCAGAACCTGATGAGATCTTGCAGCAGCACAGCTGAAATGAGAGAGAGAGAGAGAGAG 2558  
Qy 2038 GCGAGAGAAAGAAAGAAAGCGCGCGCGCGCG--CGAGGAGAGAGCTGATTTCCCGCCG 2096  
Db 2559 GCGAGAGAAAGAAAGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2618  
Qy 2097 TGGTGAAGATGAGAGATGAGAGGCGGTGCGGCGTGAAGCGGAAATGAGAGAGAGATGTGG 2156  
Db 2619 TGGTGAAGATGAGAGATGAGAGGCGGTGCGGCGTGAAGCGGAAATGAGAGAGAGATGTGG 2678  
Qy 2157 AGAGGCTGAAGCCTTACATGCTCTTGGAAATGAGGTGCCAGAGGGGAATGCAGTGCGC 2216  
Db 2679 AGAGGCTGA----- 2688  
Qy 2217 CAGCCTGTGACAAACAGCTCAGACACCGAGAGATCCCTCTTCTTCACTGAGGCGG 2276  
Db 2689 -AGCCTGTGACAAACAGCTCAGACACCGAGAGATCCCTCTTCTTCACTGAGGCGG 2747  
Qy 2277 CCAAGGACACAGGGGAGAAATGGGCGCAAGCCCGCAGCCACCTGGGCGCGCAG--GGGCGA 2335  
Db 2748 CCAAGGACACAGGGGAGAAATGGGCGCAAGCCCGCAGCCACCTTGGGCGCGCAGCGGGCGCA 2807



Db 4965 GAGATGTGATGGCCGACGCCCGGACCTGGAACGTGCTGTCTACGAGGAGAGCCCTGAAG 5024  
Qy 4531 CCGGCACGAGGACCGGCAGAGAGCTCGGAGGAGCTCAATTGCGCGGCGCCCGCTATTGT 4590  
Db 5025 CCGGCACGAGGACCGGCAGAGAGCTCGGAGGAGCTCAATTGCGCGGCGCCCGCTATTGT 5084  
Qy 4591 GCCTGAGCTGGGTAAAGCCGCGCAGAGCCCTGTGACTTATGAGAACCAAGGCGCAACCTT 4650  
Db 5085 GCCTGAGCTGGGTAAAGCCGCGCAGAGCCCTGTGACTTATGAGAACCAAGGCGCAACCTT 5144  
Qy 4651 TGGCCGCGCAGCTCCGACGAGAGTTCCGCGGTGACATATGCGGAGGCCACGCGCGCTGTA 4710  
Db 5145 TGCCCGCGCAGCTCCGACGAGAGTTCCGCGGTGACATATGCGGAGGCCACGCGCGCTGTA 5204  
Qy 4711 GAGAGGCGAGCCCTTGTGTCCAGAGAGCATCCGAGACCGAAAGCTGACGTGACGCTCG 4770  
Db 5205 GAGAGGCGAGCCCTTGTGTGTCCAGAGAGCATCCGAGACCGAAAGCTGACGTGACGCTCG 5264  
Qy 4771 TGAGATCGCCAAAGTCCCGCAGACGACCGGTGCTCGAGACCAACCAACCCCATCTCGCC 4830  
Db 5265 TGAGATCGCCAAAGTCCCGCAGACGACCGGTGCTCGAGACCAACCAACCCCATCTCGCC 5324  
Qy 4831 CTATAGAGACCTGTCTTTCGCGGCGGTGAGTGGCGTGAACCTGTATTCGCAAGCCATCCCTT 4890  
Db 5325 CTATAGAGACCTGTCTTTCGCGGCGGTGAGTGGCGTGAACCTGTATTCGCAAGCCATCCCTT 5384  
Qy 4891 GGCCTTTCGACCCACCTCCATATACCCCGCGCATCCCTCTGAGACGAGCGCTGCTACTTA 4950  
Db 5385 GGCCTTTCGACCCACCTCCATATACCCCGCGCATCCCTCTGAGACGAGCGCTGCTACTTA 5444  
Qy 4951 CCTGCCCCGACACTGTGCCCCCAACCCCACTTACCCGACCTGTACCCACCTTACTCAT 5010  
Db 5445 CCTGCCCCGACACTGTGCCCCCAACCCCACTTACCCGACCTGTACCCACCTTACTCAT 5504  
Qy 5011 CCGGCGCTAACCCCGACAGCGCGCGCTGTGAGAACGCGGACCAATCATCAATGACTCAT 5070  
Db 5505 CCGGCGCTAACCCCGACAGCGCGCGCTGTGAGAACGCGGACCAATCATCAATGACTCAT 5564  
Qy 5071 CACCTTCGAGCAGATGACCAACAACGCGCACCGCATGTGCCAGAGAGCTGATATGCT 5130  
Db 5565 CACCTTCGAGCAGATGACCAACAACGCGCGCACCGCATGTGCCAGAGAGCTGATATGCT 5624  
Qy 5131 GAGGGGCTCTTGGCCCCGAGTCTCTCGTGGCACTCAACTAGCTGCGGATCCCGAGG 5190  
Db 5625 GAGGGGCTCTTGGCCCCGAGTCTCTCGTGGCACTCAACTAGCTGCGGATCCCGAGG 5684  
Qy 5191 CATCATTCGACCTGTCCCAAGTGCACACCTGTGCTGTGCTGTGCCCCGACACAGGAC 5250  
Db 5685 CATCATTCGACCTGTCCCAAGTGCACACCTGTGCTGTGCTGTGCCCCGACACAGGAC 5744  
Qy 5251 CCAAGCCACCGGCATGAGACCGGCTTGTCTACTCCCAACCGGCGCCAGCCCTTGACAG 5310  
Db 5745 CCAAGCCACCGGCATGAGACCGGCTTGTCTACTCCCAACCGGCGCCAGCCCTTGACAG 5804  
Qy 5311 CCGCCACAGCAGCTTCCCACTCTCCCAAGAGTCCAAACAATTGACAAAACAACAC 5370  
Db 5805 CCGCCACAGCAGCTTCCCACTCTCCCAAGAGTCCAAACAATTGACAAAACAACAC 5864  
Qy 5371 CACGCTCTGTCCGAGCGGAGCGGACCGGAGATGAGAGCGGAGACCGGAGATCGGAGAG 5430  
Db 5865 CACGCTCTGTCCGAGCGGAGCGGACCGGAGATGAGAGCGGAGACCGGAGATCGGAGAG 5924  
Qy 5431 GAAAAAGTTCATCTCAAGTCCACAGCAGCGGTGAGCAGCAGCATCTGAGAGCTTGG 5490  
Db 5925 GAAAAAGTTCATCTCAAGTCCACAGCAGCGGTGAGAGCAGCATCTGAGAGCTTGG 5984  
Qy 5491 TACAGAGCAGAGCAGCGGACAGCAGCGGAGCAGCGCGGAGGTGAGAGCAGCAGCAG 5550  
Db 5985 TACAGAGCAGAGCAGCGGACAGCAGCGGAGCAGCGCGGAGGTGAGAGCAGCAGCAG 6044  
Qy 5551 CCGCGCTCTCCACTCTCCATGCCCCACAGACCTGCGCCACTCTCCCTCGAGCCCAAGATGC 5610  
Db 6045 CCGCGCTCTCCACTCTCCATGCCCCACAGACCTGCGCCACTCTCCCTCGAGCCCAAGATGC 6104

Qy 5611 CTTCCAGCAGAGACCCAGTGTGCTTCAACAACAGGACATGAAGGTATATCATCCGCTGT 5670  
Db 6105 CTTCCAGCAGAGACCCAGTGTGCTTCAACAACAGGACATGAAGGTATATCATCCGCTGT 6164  
Qy 5671 GAGAGCCGACAGCCCTCAGGTTCCTG-----AGTTCACCTCTCACCCTCTCACCCGT 5721  
Db 6165 GAGAGCCGACAGCCCTCAGGTTCCTGAGGTGAGGCTCAGGTTCACCTCTCACCCTCTCACCCGT 6224  
Qy 5722 TGGCCAGCTGCACTTCCACCTGAGCCACCACTGAGGAGGCCACCTGAGGAGGCCCTGATAG 5781  
Db 6225 TGGCCAGCTGCACTTCCACCTGAGCCACCACTGAGGAGGCCACCTGAGGAGGCCCTGATAG 6284  
Qy 5782 GGTCTACCTTACCTCATGAGAGCCGCTGTGTGCTGCTCCAGAGAGGCCCGCGGTGCGCG 5841  
Db 6285 GGTCTACCTTACCTCATGAGAGCCGCTGTGTGCTGCTCCAGAGAGGCCCGCGGTGCGCG 6344  
Qy 5842 GCCAGAGCGGCCCGGACAGACATCGGACATGCTTCTCTGCAAGGCCCGCCAGCCGCTC 5901  
Db 6345 GCCAGAGCGGCCCGGACAGACATCGGACATGCTTCTCTGCAAGGCCCGCCAGCCGCTC 6404  
Qy 5902 CCGGCTGAGAGCCGCTCTCCCGCAGAGAGGCTCGAGGCCCGGAGCCCTTATGAGCTCC 5961  
Db 6405 CCGGCTGAGAGCCGCTCTCCCGCAGAGAGGCTCGAGAGGCCCGGAGCCCTTATGAGCTCC 6464  
Qy 5962 TGTCTCTGAGCAGCGCACCATGCGCCGACCCCTGCGAAGAACTTCGACCTCAACACGC 6021  
Db 6465 TGTCTCTGAGCAGCGCACCATGCGCCGACCCCTGCGAAGAACTTCGACCTCAACACGC 6524  
Qy 6022 CAGCCCGGACCCGCGCGGCGCACCTGTGCTCGGCTTCGAGCCCGGACCGGAGAAAGCTCA 6081  
Db 6525 CAGCCCGGACCCGCGCGGCGCACCTGTGCTCGGCTTCGAGCCCGGACCGGAGAAAGCTCA 6584  
Qy 6082 AAGTAAACCTTTTTCATTCAGAGAACTGGAATCTCGTCTGAGGTTTACACGCGAGAG 6141  
Db 6585 AAGTAAACCTTTTTCATTCAGAGAACTGGAATCTCGTCTGAGGTTTACACGCGAGAG 6644  
Qy 6142 CTACAGCCCGGAGGAGGTGAGAGCCGCTGAGCCCTGTGAGCTCACCCAGTTCGACCAAGA 6201  
Db 6645 CTACAGCCCGGAGGAGGTGAGAGCCGCTGAGCCCTGTGAGCTCACCCAGTTCGACCAAGA 6704  
Qy 6202 CAGAGGCTCTCCCAAGACCTGTGAAGAGCTGACAAAGGCCACTGTGAAGGAGAGCTGCG 6261  
Db 6705 CAGAGGCTCTCCCAAGACCTGTGAAGAGCTGACAAAGGCCACTGTGAAGGAGAGCTGCG 6764  
Qy 6262 GCCCAAGCAGGAGCCCGGTGAAGCTTGTGCGGAGAGGCGGCCACCTCCCAACCTTGC 6321  
Db 6765 GCCCAAGCAGGAGCCCGGTGAAGCTTGTGCGGAGAGGCGGCCACCTCCCAACCTTGC 6824  
Qy 6322 GCCGCTGCTGAGAGCCAGCCCTGTGTCAGCCCGCTGTGTCAGAACCGGCCCAAGGAGTCAA 6381  
Db 6825 GCCGCTGCTGAGAGCCAGCCCTGTGTCAGCCCGCTGTGTCAGAACCGGCCCAAGGAGTCAA 6884  
Qy 6382 AGGTACCAAGGAGGTGTACCTGTGGCCAGACATCATGTAGGTATATCAACAGAGCTTA 6441  
Db 6885 AGGTACCAAGGAGGTGTACCTGTGGCCAGACATCATGTAGGTATATCAACAGAGCTTA 6944  
Qy 6442 CACCCGAGCACAACCAAGCAGCTCAGCGCAACCCCTGCGCGGCCCTCTACTCTCTTCC 6501  
Db 6945 CACCCGAGCACAACCAAGCAGCTCAGCGCAACCCCTGCGCGGCCCTCTACTCTCTTCC 7004  
Qy 6502 TGGGCGCAGCTGCCCGCTCTTGAACCTTCGCGCGCCACCCAGTGAACCTTACTCCGCG 6561  
Db 7005 TGGGCGCAGCTGCCCGCTCTTGAACCTTCGCGCGCCACCCAGTGAACCTTACTCCGCG 7064  
Qy 6562 CCGGAGCATATGTGCTCCGCGCTGTGCTTCCCGCCACAGCAAGGAGGAGAGGTCTCC 6621  
Db 7065 CCGGAGCATATGTGCTCCGCGCTGTGCTTCCCGCCACAGCAAGGAGGAGAGGTCTCC 7124  
Qy 6622 AGAGCAAAACAAGACGTGTGCTTGTGAGTGTGAGAGACGTATTGAACCTGTGTCCC 6681  
Db 7125 AGAGCAAAACAAGACGTGTGCTTGTGAGTGTGAGAGACGTATTGAACCTGTGTCCC 7184



QY 6682 ACCGAGGCGATGACGAGCCGAGGCACTCCCGAGTGTGTGTACCCGCTGTGATCCG 6741  
DB ACCGAGGCGATGACGAGGCGAGGCACTCCCGAGTGTGTGTACCCGCTGTGATCCG 7244  
QY 6742 GGTATGGGGAAACAGACGAGCCGACGAGATGGCTTCCAGTCTTCAGGCAACACAGCC 6801  
DB GGTATGGGGAAACAGACGAGCCGACGAGATGGCTTCCAGTCTTCAGGCAACACAGCC 7304  
QY 6802 GCGGCGAGCTTCTTGAAGAGTGAACCGAGAGCACTCCGCAATGGTCAAGTCAAGAA 6861  
DB GCGGCGAGCTTCTTGAAGAGTGAACCGAGAGCACTCCGCAATGGTCAAGTCAAGAA 7364  
QY 6862 GCAAGAGATCAACAAAGCTGAACCCCAACCGGAATGAGCTGAATATCAATATCAG 6921  
DB GCAAGAGATCAACAAAGCTGAACCCCAACCGGAATGAGCTGAATATCAATATCAG 7424  
QY 6922 CCAGCTTGGAGCGAGATCTTCAATATGCGCCATATCAAGGAAACAGGCTTATGACTTA 6981  
DB CCAGCTTGGAGCGAGATCTTCAATATGCGCCATATCAAGGAAACAGGCTTATGACTTA 7484  
QY 6982 TGAAGACCGAGCGGTGACAGGAATGCGAGCAACAAATGGGGCTGAGGSCATATTTAG 7041  
DB TGAAGACCGAGCGGTGACAGGAATGCGAGCAACAAATGGGGCTGAGGSCATATTTAG 7544  
QY 7042 AAAGGCACTCATGGGTAAATATGACAGTGGAGAGAGTCCCGCCGCTCAGCGCAATGC 7101  
DB AAAGGCACTCATGGGTAAATATGACAGTGGAGAGAGTCCCGCCGCTCAGCGCAATGC 7604  
QY 7102 TTTTAAACCTTGATGCGAGTGCAGGCTGCGCTGCTATGSCCATTAACGCTGTGTA 7161  
DB TTTTAAACCTTGATGCGAGTGCAGGCTGCGCTGCTATGSCCATTAACGCTGTGTA 7664  
QY 7162 CGGACGAGATGACACACACTGACCTGCGAGGTGGCGGGGAAAGGCCAAAGTCTCTG 7221  
DB CGGACGAGATGACACACACTGACCTGCGAGGTGGCGGGGAAAGGCCAAAGTCTCTG 7724  
QY 7222 CAGACCCAGACCGGAAAGCAAGTCCCGGCGCCGCGCTGAGCATTTGGGGACCGGCG 7281  
DB CAGACCCAGACCGGAAAGCAAGTCCCGGCGCCGCGCTGAGCATTTGGGGACCGGCG 7784  
QY 7282 ACCCTGTGTCTCTCAAGTGCATGCGAGGAGACTGCAACCGCGGACCGCTGTCAACAA 7341  
DB ACCCTGTGTCTCTCAAGTGCATGCGAGGAGACTGCAACCGCGGACCGCTGTCAACAA 7844  
QY 7342 CGGCGTGTGGAGAGACAGGCGCTGCTCGGAGGTTCCAGGCCATTTCCCTTAACAACCCCT 7401  
DB CGGCGTGTGGAGAGACAGGCGCTGCTCGGAGGTTCCAGGCCATTTCCCTTAACAACCCCT 7904  
QY 7402 GATCATGCGGCTGACAGCGGAGTGTCAATGCTTCCCAACCCCAACCGGCTTCCCGCGG 7461  
DB GATCATGCGGCTGACAGCGGAGTGTCAATGCTTCCCAACCCCAACCGGCTTCCCGCGG 7964  
QY 7462 CAGCGGAGCCCTGCTGAGCGCCCAACCGCTTGGAGAGAGAGCCCAAGCACTGTCTTG 7521  
DB CAGCGGAGCCCTGCTGAGCGCCCAACCGCTTGGAGAGAGAGCCCAAGCACTGTCTTG 8024  
QY 7522 CTTCGAGTACAGACACTCTCCGACAGGAGTGAATCAAGACGCGGGGGGGGGGGGG 7581  
DB CTTCGAGTACAGACACTCTCCGACAGGAGTGAATCAAGACGCGGGGGGGGGGGGGGG 8084  
QY 7582 CGGTUBERSPERLUGTCAAGTCCAGCGAGCAACAGGAACGCGCTTGCAGAGAGCGGAGCG 7641  
DB CGGTUBERSPERLUGTCAAGTCCAGCGAGCAACAGGAACGCGCTTGCAGAGAGCGGAGCG 8135  
QY 7642 GCTGCGGACTCCCCCAACCAAGAGAGAGCCCTGAGTCCGCTGCGCTCATTCATCT 7701  
DB GCTGCGGACTCCCCCAACCAAGAGAGAGCCCTGAGTCCGCTGCGCTCATTCATCT 8195  
QY 7702 GTTCGCTCAGAGCGGAGTCTTGTGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 7761  
DB GTTCGCTCAGAGCGGAGTCTTGTGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 8255  
QY 7762 CTGGCCCTGTGAGACCTTACTCAGGGAGATGTTTACTGTGTGCTGGAGAGGAGGAGAA 7821

DB 8256 CTGGCCCTGTGAGACCTTACTCAGGGAGATGTTTACTGTGTGCTCGGAAAGGAGGAGAA 8315  
QY 7822 GGGGCGGGGAGGGGGGACGCGAGGCGTGTGAGAGCAACAGAGCGGCGAGGGGGCC 7881  
DB GGGGCGGGGAGGGGGGACGCGAGGCGTGTGAGAGCAACAGAGCGGCGAGGGGGCC 8375  
QY 7882 AGGAGCCAAAGAGAGATGACACGACCTCAAGCAGCTCTCCCGAATGCAATTTG 7941  
DB AGGAGCCAAAGAGAGATGACACGACCTCAAGCAGCTCTCCCGAATGCAATTTG 8435  
QY 7942 GAAACCAAGCTTAACTGAGCTGCAAGCCCGCGCTTCCCTGCGCTCATTCGCT 8001  
DB GAAACCAAGCTTAACTGAGCTGCAAGCCCGCGCTTCCCTGCGCTCATTCGCT 8495  
QY 8002 TTAGGCTCTGAGCAGATGAGACGAGGCGCTGTCAAGCCCGAGTGGCTGCTCGGCTC 8061  
DB TTAGGCTCTGAGCAGATGAGACGAGGCGCTGTCAAGCCCGAGTGGCTGCTCGGCTC 8555  
QY 8062 CCACAGACTGCGCCAGCCCAACAGATTTGCTGAAACCAAGTCAAGCCAGGTGGCGAGCA 8121  
DB CCACAGACTGCGCCAGCCCAACAGATTTGCTGAAACCAAGTCAAGCCAGGTGGCGAGCA 8615  
QY 8122 AAAGGCGCAGGTGCGGCTGAGGAGAAAGGAGTCCGAGAGCTGACTGTTTTTTTAC 8181  
DB AAAGGCGCAGGTGCGGCTGAGGAGAAAGGAGTCCGAGAGCTGACTGTTTTTTTAC 8675  
QY 8182 ACATCGTTCGCGAGGCGTGGAGAAAGGAGAGTAAATGATGTGTGTTTCAAG 8241  
DB ACATCGTTCGCGAGGCGTGGAGAAAGGAGAGTAAATGATGTGTGTTTCAAG 8735  
QY 8242 GTATATTTTGAATACCTTCAATGAATTAATTCAGATGTTTACGCAAGAGAGACTTAC 8301  
DB GTATATTTTGAATACCTTCAATGAATTAATTCAGATGTTTACGCAAGAGAGACTTAC 8795  
QY 8302 CAGTATTTACTGCTGTGTGCTTGTGATCTCTGTTAACGTTCAAGAGCGTGTGACGCC 8361  
DB CAGTATTTACTGCTGTGTGCTTGTGATCTCTGTTAACGTTCAAGAGCGTGTGACGCC 8855  
QY 8362 GACAGTCGAGTACCCCATCATCTGACAGACCAAGGGGGGGGGGACGCTGCG -TCAAGGCC 8420  
DB GACAGTCGAGTACCCCATCATCTGACAGACCAAGGGGGGGGGGACGCTGCGCTGCG -TCAAGGCC 8915  
QY 8421 CGCTGTGCTCTCCCTCCCTCCCTTCTTGGGAGAAATGATGATGCGTATCTGTGCG 8480  
DB CGCTGTGCTCTCCCTCCCTCCCTTCTTGGGAGAAATGATGATGCGTATCTGTGCG 8975  
QY 8481 CGCATTTCGAGAGGAGTGTGATTTCTGTCAATTTACACAGTGTGTTTAAATTAAGC 8540  
DB CGCATTTCGAGAGGAGTGTGATTTCTGTCAATTTACACAGTGTGTTTAAATTAAGC 9035  
QY 8541 GAATTTACTCCAAAAA 8558  
DB GAATTTACTCCAAAAA 9053

RESULT 12  
ADQ23294/c  
ID ADQ23294 standard; DNA; 9079 BP.

AC ADQ23294;  
XX  
DT 26-AUG-2004 (first entry)

XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 6114.  
XX  
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;  
XX de.  
XX  
XX Homo sapiens.  
XX  
XX PN MO2004048938-A2.



QY 1620 ACGAAGAGAAAGCTCTTCAAGAGAGAGACAGACGACCTCAAGGGAGAGACAACGACG 1679  
DB 6939 ACGAAGAGAAAGACTCTTCAAGAGAGAGACAGACGACCTCAAGGGAGAGACAACGACG 6880  
QY 1680 AGAAGAGAGCTGTGGCTCCAAAGGCGCGAAATCTGCAACAGCCAGAGAAAGACGCAAG 1739  
DB 6879 AGAAGAGAGCTGTGGCTCCAAAGGCGCGAAATCTGCAACAGCCAGGGAGAGACGCAAG 6820  
QY 1740 GCCGACACCCGCTCAATGGCTAATGAGGCAACGCGAGAGGGCCATGACCCCGACG 1799  
DB 6819 GCCGACACCCGCTCAATGGCTAATGAGGCAACGCGAGAGGGCCATGACCCCGACG 6760  
QY 1800 AAGAGCCGAGCTGTGGCTCCATGAGCTGATGAGATTCTCGCTGAGACAGAAAGAA 1859  
DB 6759 AAGAGCCGAGCTGTGGCTCCATGAGCTGATGAGATTCTCGCTGAGACAGAAAGAA 6700  
QY 1860 TGGAAACAGCCAAAGAAAGTCTTCTGGAACACGCGCGCACTGGTGCCATGCGCCGGA 1919  
DB 6699 TGGAAACAGCCAAAGAAAGTCTTCTGGAACACGCGCGCACTGGTGCCATGCGCCGGA 6640  
QY 1920 TGGTGGGCTCCAAAGCTGTGTGCGAGGTGAAGACTTCTTCACTCAAGAAAGAGC 1979  
DB 6639 TGGTGGGCTCCAAAGCTGTGTGCGAGGTGAAGACTTCTTCACTCAAGAAAGAGC 6580  
QY 1980 AGAACCTCGATGAGATCTTGCAGCAGACACAGCTGGAAGTGAAGAGAGAGAGCGCG 2039  
DB 6579 AGAACCTCGATGAGATCTTGCAGCAGACACAGCTGGAAGTGAAGAGAGAGAGCGCG 6520  
QY 2040 GGAAGAAAGAAAGAAAGCGCGCGCGCGCGCGCAAGAGAGCTGCAATTCGCGCGTG 2099  
DB 6519 GGAAGAAAGAAAGAAAGCGCGCGCGCGCGCGCAAGAGAGCTGCAATTCGCGCGTG 6460  
QY 2100 TGGAGGATGAGAGATGAGAGGTGTGGGCTGAGCGGAAATGAGAGAGAGATGTGTGAGG 2159  
DB 6459 TGGAGGATGAGAGATGAGAGGTGTGGGCTGAGCGGAAATGAGAGAGAGATGTGTGAGG 6400  
QY 2160 AAGCTGAAGCTTTCACATGCTCTGAGATGAGGTGCCAGAGGGAGATCAGTGCGCCAG 2219  
DB 6399 AAGCTGA-----AG 6391  
QY 2220 CCACTGTCAACACAGCTCAAGACACCGAGAGCATCCCTCTCTCTCACTAGAGCGCGCA 2279  
DB 6390 CCACTGTCAACACAGCTCAAGACACCGAGAGCATCCCTCTCTCTCACTAGAGCGCGCA 6331  
QY 2280 AAGACAAGGGGAGAAATGGGCCCCAAGCCCCAGCCACCTGTGGGCGCGGACGCGCC 2339  
DB 6330 AAGACAAGGGGAGAAATGGGCCCCAAGCCCCAGCCACCTGTGGGCGCGGACGCGCC 6271  
QY 2340 CAGGCCCAACCCCAACCGAGAGGACATCCCGGGCCCCATTGAGCCCAACCCGGGCT 2399  
DB 6270 CAGGCCCAACCCCAACCGAGAGGACATCCCGGGCCCCATTGAGTCAACCCGGGCT 6211  
QY 2400 CTGAAGCCACCGGAGCCCTTACGCCCCCAACAGACCCCATTCGCTTGCACCTCTC 2459  
DB 6210 CTGAAGCCACCTTACGCCCCCAACAGACCCCATTCCTTTCACCTCTC 6151  
QY 2460 CTGTGTCTCCCAAGAGAGAGAGAGAGAGACCGACACGCGCCCCCAAGTGAAGAG 2519  
DB 6150 CTGTGTCTCCCAAGAGAGAGAGAGAGAGACCGACACGCGCCCCCAAGTGAAGAG 6091  
QY 2520 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2579  
DB 6090 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6031  
QY 2580 AGCCCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2639  
DB 6030 AGCCCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5971  
QY 2640 CGAGAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2699  
DB 5970 CGAGAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5911

QY 2700 GCGGAGAGGCACTACTGCCAAGAGCTCGGGAGCCCCCGCCAGAGACAGAGACTCCAGTGA 2759  
DB 5910 GCGGAGAGGCACTACTGCCAAGAGCTCGGGAGCCCCCGCCAGAGACAGAGACTCCAGTGA 5851  
QY 2760 CTTGAGTGCAGACAGAGTGAATGAGGCGGAGCGCGCAAGAAACCGGCTGTGTCC 2819  
DB 5850 CTTGAGTGCAGACAGAGTGAATGAGGCGGAGCGCGCAAGAAACCGGCTGTGTCC 5791  
QY 2820 CAAGGCCAGCTCTCAACCCGACTGGGAGACCCCGGGCAATGACTCAACCCAGAGAG 2879  
DB 5790 CAAGGCCAGCTCTCAACCCGACTGGGAGACCCCGGGCAATGACTCAACCCAGAGAG 5731  
QY 2880 CACTGACCTGAAGCAGCTGAACAGACAGCGGCTGCATCCCGCCATCCAGTCAACA 2939  
DB 5730 CACTGACCTGAAGCAGCTGAACAGACAGCGGCTGCATCCCGCCATCCAGTCAACA 5671  
QY 2940 AAGTCAATGAGCCCCCGGAGAGAGAGAGCTTCCACCAAGCTCTCCCGAGCCAC 2999  
DB 5670 AAGTCAATGAGCCCCCGGAGAGAGAGAGCTTCCACCAAGCTCTCCCGAGCCAC 5611  
QY 3000 CGCACCGGAAAACTGCAAGCGGAGAGAGAGCGCCCTCAGACGCTGGAGAGAGCGCC 3059  
DB 5610 CGCACCGGAAAACTGCAAGCGGAGAGAGAGCGCCCTCAGACGCTGGAGAGAGCGCC 5551  
QY 3060 GGGGCAAGAGCAGAGAGCGCGGACCCCGCGCAAGAG----- 3100  
DB 5550 GGGGCAAGAGCAGAGAGCGCGGACCCCGCGCAAGAGAGCGCTGTCTC 5491  
QY 3101 -----GCTTCCGAGCGAGGCGCCAGAGCTGCTGGGAGACCCCTTGTGTGACTTCCG 3155  
DB 5490 TCCAGACCTTCGAGCGAGCGCCAGAGAGCTGCTGGGAGACCCCTTGTGTGACTTCCG 5431  
QY 3156 GCTGCGCTTCCCGGAGCGCCCGGAGAGTCAAGGCTTCCCGAGATGCCCGGAGC 3215  
DB 5430 GCTGCGCTTCCCGGAGCGCCCGGAGAGTCAAGGCTTCCCGAGATGCCCGGAGC 5371  
QY 3216 CCTCAAGCTTCTCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 3275  
DB 5370 CCTCAAGCTTCTCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTT 5311  
QY 3276 CCGGAGCGCTTCTGCGGCGCCCAACCAATCTCAACCGGCTCTCTCATCTCTCTG 3335  
DB 5310 CCGGAGCGCTTCTGCGGCGCCCAACCAATCTCAACCGGCTCTCTCATCTCTCTG 5251  
QY 3336 CCAAGACCCCAAGCGCTTCTGAGAGCAATAGAGCCATCTCCAGAGAAATCTCGTCC 3395  
DB 5250 CCAAGACCCCAAGCGCTTCTGAGAGCAATAGAGCCATCTCTCCAGAGAAATCTCGTCC 5191  
QY 3396 AGCTCCAGCTCCGTACTCAGAGCATGCAAGAGCCCGGCTGGGCTCTCAACATGGGCG 3455  
DB 5190 AGCTCCAGCTCCGTACTCAGAGCATGCAAGAGCCCGGCTGGGCTCTCAACATGGGCG 5131  
QY 3456 TGCCTTGCATGAGACCCCAAAAGCTGCACTTCAAGCGAGTGAAGAGAGAGAGAG 3515  
DB 5130 TGCCTTGCATGAGACCCCAAAAGCTGCACTTCAAGCGAGTGAAGAGAGAGAGAG 5071  
QY 3516 TGTCCCAAGGGGCAAGGCTGGGCGCAACCGGAGAGCTTGGGGTGCCTCAAGCGAGAG 3575  
DB 5070 TGTCCCAAGGGGCAAGGCTGGGCGCAACCGGAGAGCTTGGGGTGCCTCAAGCGAGAG 5011  
QY 3576 GGTCCGTGCTGAGAGAGAGAGCTTGGGCTCAGTTCCGGGCGGAGAGATCAACAAAGAG 3635  
DB 5010 GGTCCGTGCTGAGAGAGAGAGCTTGGGCTCAGTTCCGGGCGGAGAGATCAACAAAGAG 4951  
QY 3636 TTCCAGACACAGGCTGCTTGCACAGCGCATCAATACCGGCTCATACCCACG 3695  
DB 4950 TTCCAGACACAGGCTGCTTGCACAGCGCATCAATACCGGCTCATACCCACG 4891  
QY 3696 GGAAGCGAGCTGAGCTTCCTGTAACAAGGCGACATCAAGAGATCAATCCGCGAGAGAG 3755  
DB 4890 GGAAGCGAGCTGAGCTTCCTGTAACAAGGCGACATCAAGAGATCAATCCGCGAGAGAG 4831  
QY 3756 CGAGTCTGTTGACCGCGGCGGAGAGAGAGCTGCCCAAGGCGACGTCATCTACGAAG 3815

|||||  
Db CGAGTCGCTTGGACCGCGCCCGGAGAGCAGCTCCCAAGGCGCAGCTACCTACGAAAG 4771  
Qy GCAAGAAAGGGCCAGCTTTGTCTATGAGGGGTGATGTCTGTGACCCAGTCTCCAAAG 3875  
Db GCMAAAGAGGCCAGCTTGTCTGTATGAGGGTGGCATGTCTGTGACCAAGTCTCCAAAG 4711  
Qy AGGACGGAGAAAGCAGCTCAGGACCCCCCATAGAGCGGCGGCCCAAGCGCACCTATG 3935  
Db AGGACGGAGAAAGCAGCTCAGGACCCCCCATAGAGCGGCGGCCCAAGCGCACCTATG 4651  
Qy ACATATGAGAGGGCGCTGTGGCAGAGCCATCTCTCAGCCAGCATTCGAAGGTCTCATGG 3995  
Db ACATATGAGAGGGCGCTGTGGCAGAGCCATCTCTCAGCCAGCATTCGAAGGTCTCATGG 4591  
Qy GCCCGGCATCCCGCGGAGCGACAGGCCCCACACACTCAAAAGAGCAGCACATCC 4055  
Db GCCGTGCATCCCGCGGAGCGACAGGCCCCACACACTCAAAAGAGCAGCACATCC 4531  
Qy GCGGCTCATCAACAAGGAGATCCCTCGGTCTTACGTGAGGACAGAGAGCTACCTGC 4115  
Db GCGGCTCATCAACAAGGAGATCCCTCGGTCTTACGTGAGGACAGAGAGCTACCTGC 4471  
Qy GTCCGGAAGGCCAGCTCTTAAAGCGGAGGGCAGCGCTCCGCCCCACCGCCCTCACGG 4175  
Db GTCCGGAAGGCCAGCTCTTAAAGCGGAGGGCAGCGCTCCGCCCCACCGCCCTCACGG 4411  
Qy ACTGACCGAGGCGCTAACAAAGCGAGGCCCTGGGGCCCCCTGAAGCTGAAGCGGCGCATG 4235  
Db ACTGACCGAGGCGCTAACAAAGCGAGGCCCTGGGGCCCCCTGAAGCTGAAGCGGCGCATG 4351  
Qy AGGGCTGTGGCCACCGGTGAAGGAGGGGGCGGCTCATTCATAGAGATCCGGCGAGG 4295  
Db AGGGCTGTGGCCACCGGTGAAGGAGGGGGCGGCTCATTCATAGAGATCCGGCGAGG 4291  
Qy AGCTGCGGACACCGCCGAGCTGCCCTGGCCCGCGGCGCTCAAGGAGGCTCATCA 4355  
Db AGCTGCGGACACCGCCGAGCTGCCCTGGCCCGCGGCGCTCAAGGAGGCTCATCA 4231  
Qy CGCAGGGGACCCCGCTCAAGTACGACACCGGGGCTCCACATGAGCTCCAAAAGCACG 4415  
Db CGCAGGGGACCCCGCTCAAGTACGACACCGGGGCTCCACATGAGCTCCAAAAGCACG 4171  
Qy AGCTACGCTCTCTCATCGGCAAGCCCCCGGCGGACGTTCCACCCGTGACCTGTGATG 4475  
Db AGCTACGCTCTCTCATCGGCAAGCCCCCGGCGGACGTTCCACCCGTGACCTGTGATG 4111  
Qy TGATGGCCGAGCGCCCGGCGACTGTGAACGTGCTGTACAGAGAGAGCTGAAGACCGGC 4535  
Db TGATGGCCGAGCGCGCGGCGACTGTGAACGTGCTGTACAGAGAGAGCTGAAGACCGGC 4051  
Qy CAGGAGCGGCGAGGCTCGGGGGGCTCATTTGGCGGGGCGCCCGGTCAATTGGGCGT 4595  
Db CAGGAGCGGCGAGGCTCGGGGGGCTCATTTGGCGGGGCGCCCGGTCAATTGGGCGT 3991  
Qy AGCTGGTAAAGCGCGGAGAGCCCCCTTGACCTTATGAGGACCAAGGAGACCTTTGGCG 4655  
Db AGCTGGTAAAGCGCGGAGAGCCCCCTTGACCTTATGAGGACCAAGGAGACCTTTGGCG 3931  
Qy GCCACTCTCCAGAGGTTCCCGGTGACCATGCGGAGGCCACCGCGGCGCTGTGAGAGG 4715  
Db GCCACTCTCCAGAGGTTCCCGGTGACCATGCGGAGGCCACCGCGGCGCTGTGAGAGG 3871  
Qy GCAGCTTTTCGTCAGCAAGGAGATCCAGGAGCCGAAGCTGACGTGACGCTCGTGAAG 4775  
Db GCAGCTTTTCGTCAGCAAGGAGATCCAGGAGCCGAAGCTGACGTGACGCTCGTGAAG 3811  
Qy TCGCCAGTCTCCCGCAGACCGGTGCCGAGCACCAACCCCACTTCCTGCGCTATG 4835  
Db TCGCCAGTCTCCCGCAGACCGGTGCCGAGCACCAACCCCACTTCCTGCGCTATG 3751  
Qy AGCACCCTGCTTGGGGCGTGAATGCGGTGACCTGTATCGCAGCCACATCCCCCTGGCCT 4895  
|||||

Db AGCACCCTGCTTGGGGCGTGAATGCGGTGACCTGTATCGCAGCCACATCCCCCTGGCCT 3691  
Qy TCGACCCCACTTCATACCCCGGACATCCCTTGGAGCGAGCGGCTCTTACTTGC 4955  
Db TCGACCCCACTTCATACCCCGGACATCCCTTGGAGCGAGCGGCTCTTACTTGC 3631  
Qy CCGGACCTTGGCCCCCAACCCCACTTACCCGACCTGTACCCACCTTACTTCCGG 5015  
Db CCGGACCTTGGCCCCCAACCCCACTTACCCGACCTGTACCCACCTTACTTCCGG 3571  
Qy GCTACCCCGACACGGCGGGCGGTGAGAAACGGGACCAATCATATGATGATCATCACT 5075  
Db GCTACCCCGACACGGCGGGCGGTGAGAAACGGGACCAATCATATGATGATCATCACT 3511  
Qy CGCAGCAGATGACCAACAACGAGCCACCGCATGGCCCAAGAGCTGATATGTAGAGG 5135  
Db CGCAGCAGATGACCAACAACGAGCCACCGCATGGCCCAAGAGCTGATATGTAGAGG 3451  
Qy GCTTCTGCCCCCGAGTCTTGTGGCATCTCACTACGCTGCGGGTCCCGAGGACATCA 5195  
Db GCTTCTGCCCCCGAGTCTTGTGGCATCTCACTACGCTGCGGGTCCCGAGGACATCA 3391  
Qy TCGACCTGTCCCAAGTGGCACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5255  
Db TCGACCTGTCCCAAGTGGCACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3331  
Qy CCACCGCATGGAACCGCTTGTCTTACCTCCACCGGCGCCAGGCTTCAGACACCGCC 5315  
Db CCACCGCATGGAACCGCTTGTCTTACCTCCACCGGCGCCAGGCTTCAGACACCGCC 3271  
Qy ACAAGAGTCTCCCACTCTCCCAAGAGGTCCAAACAATTGACAAACCAACCAACAGT 5375  
Db ACAAGAGTCTCCCACTCTCCCAAGAGGTCCAAACAATTGACAAACCAACCAACAGT 3211  
Qy CCGTCCGAGCGGAGGAGACCGGGATCGAGCGGGGACCGGGATGCGGGAGCGGGAAA 5435  
Db CCGTCCGAGCGGAGGAGACCGGGATCGAGCGGGGACCGGGATGCGGGAGCGGGAAA 3151  
Qy AGTCCATCTCAACGTCCACCAAGAGGTCCAAACAATTGACAAACCAACCAACAGT 5495  
Db AGTCCATCTCAACGTCCACCAAGAGGTCCAAACAATTGACAAACCAACCAACAGT 3091  
Qy AGCAGAGAGCGGCGAGCGGACGCGGCGGAGGAGCGAGCCCATCTGTGAGACCTGTGACG 5555  
Db AGCAGAGAGCGGCGAGCGGACGCGGCGGAGGAGCGAGCCCATCTGTGAGACCTGTGACG 3031  
Qy CCTCCCATCTCCAGGCCACAGGACCTGCCCATCTCCCTCGGACCCAGGATGCGCTCC 5615  
Db CCTCCCATCTCCAGGCCACAGGACCTGCCCATCTCCCTCGGACCCAGGATGCGCTCC 2971  
Qy AGCAGAGACCAAGTGTCTTCAACAACAGGCAATGAAGGTATCATCAACCGCTGTGAGC 5675  
Db AGCAGAGACCAAGTGTCTTCAACAACAGGCAATGAAGGTATCATCAACCGCTGTGAGC 2911  
Qy CCAGCAAGCCCAAGTCTTGAAGTCCACCTTCACTCTCAACCGTTCGCCAGCTGCCA 5735  
Db CCAGCAAGCCCAAGTCTTGAAGTCCACCTTCACTCTCAACCGTTCGCCAGCTGCCA 2851  
Qy CATTTCCACCTGTGCCACCACTGTGCCCATGTGGGCGGACCTGTGATGGGTCTTACCT 5795  
Db CATTTCCACCTGTGCCACCACTGTGCCCATGTGGGCGGACCTGTGATGGGTCTTACCT 2791  
Qy TCATGAGACCGGTCTTGTGTCGCAAGAGAGCCCCCGGGGTGCGCCGAGCAGAGCGGCCC 5855  
Db TCATGAGACCGGTCTTGTGTCGCAAGAGAGCCCCCGGGGTGCGCCGAGCAGAGCGGCCC 2731  
Qy GAGCAGACACCGGCATGCTTCTGTGCAAGCCCCCAGCCGCTTCGAGGCTGTGAGCCCG 5915  
Db GAGCAGACACCGGCATGCTTCTGTGCAAGCCCCCAGCCGCTTCGAGGCTGTGAGCCCG 2671  
Qy CCTCTCTCCCCAGCAGGAGGCTGTGAGCCCCCGGCTTGTGTCTGTCTTGTGCGACG 5975  
Db CCTCTCTCCCCAGCAGGAGGCTGTGAGCCCCCGGCTTGTGTCTGTCTTGTGCGACG 2611  
|||||

QY 5976 CCAACATCGCCCGCACCCCTGCGAAGAACTCGCACTTCAACAAGCCGACGCCGCG 6035  
| | | | |  
Db 2610 CCAACATCGCCCGCACCCCTGCGAAGAACTCGCACTTCAACAAGCCGACGCCGCG 2551  
| | | | |  
QY 6036 CGGCGCCACCTGCTCGGCTCGGACCCGCAACCGGAAAAAGACTCAAGTAAACCTTTT 6095  
| | | | |  
Db 2550 CGGCGCCACCTGCTCGGCTCGGACCCGCAACCGGAAAAAGACTCAAGTAAACCTTTT 2491  
| | | | |  
QY 6096 CCATCCAGGAACCTGGAACCTCCGTTCTGCGTTAACAACGCGACACTTCAAGCCCGGAG 6155  
| | | | |  
Db 2490 CCATCCAGGAACCTGGAACCTCCGTTCTGCGTTAACAACGCGACACTTCAAGCCCGGAG 2431  
| | | | |  
QY 6156 GGGTGGAGCCCGTCAGCCCTGTGAGCTCAACAGTCTGACCCACGACAAAGGGCTCCCA 6215  
| | | | |  
Db 2430 GGGTGGAGCCCGTCAGCCCTGTGAGCTCAACAGTCTGACCCACGACAAAGGGCTCCCA 2371  
| | | | |  
QY 6216 AGCACTTGAAGAGCTCGAACAAGGCCACTTGAAGGGGAGAGTGGCGGCCCAAGCAGCCAG 6275  
| | | | |  
Db 2370 AGCACTTGAAGAGCTCGAACAAGGCCACTTGAAGGGGAGAGTGGCGGCCCAAGCAGCCAG 2311  
| | | | |  
QY 6276 GGGCCGTTGAAGCTTGGCGGGGAGAGCCGCCCACTTCCCAACCTTGGCGGCTGCTGAGA 6335  
| | | | |  
Db 2310 GGGCCGTTGAAGCTTGGCGGGGAGAGCCGCCCACTTCCCAACCTTGGCGGCTGCTGAGA 2251  
| | | | |  
QY 6336 GCGAGCCCTGCTCGACCCGCTGCTCAAGACGCGCCAGGGGTCAAAAGTCAACAGCGGG 6395  
| | | | |  
Db 2250 GCGAGCCCTGCTCGACCCGCTGCTCAAGACGCGCCAGGGGTCAAAAGTCAACAGCGGG 2191  
| | | | |  
QY 6396 TGGTCACTCTGAGCCGACATCATGATGAGTATCAACAAGACTCAACCCGCGACACC 6455  
| | | | |  
Db 2190 TGGTCACTCTGAGCCGACATCATGATGAGTATCAACAAGACTCAACCCGCGACACC 2131  
| | | | |  
QY 6456 CACAGAGCTCAGCGGACCCCTGCGCCGCCCTTCACTCTTCCCTGGGGCCAGCTGCC 6515  
| | | | |  
Db 2130 CACAGAGCTCAGCGGACCCCTGCGCCGCCCTTCACTCTTCCCTGGGGCCAGCTGCC 2071  
| | | | |  
QY 6516 CGTCTGAGACCTCGCGCGCCCAACGATGACTCTTACCTCCGCGCCCGGACCATGGTG 6575  
| | | | |  
Db 2070 CGTCTGAGACCTCGCGCGCCCAACGATGACTCTTACCTCCGCGCCCGGACCATGGTG 2011  
| | | | |  
QY 6576 CCGCGGCGCGTGGCTCCCGCAACGGAAGGGGCAAGAGTCTTCAGAGCCCAACAG 6635  
| | | | |  
Db 2010 CCGCGGCGCGTGGCTCCCGCAACGGAAGGGGCAAGAGTCTTCAGAGCCCAACAG 1951  
| | | | |  
QY 6636 CGTCCGTCTTGGGTGGTGGTGAAGAGCGTATTGAACCTGTGTCCTCCGAGGGGATGA 6695  
| | | | |  
Db 1950 CGTCCGTCTTGGGTGGTGGTGAAGAGCGTATTGAACCTGTGTCCTCCGAGGGGATGA 1891  
| | | | |  
QY 6696 CGGAGCGAGGGGACTCCGGAGTGTGTGAACCGCTGCTGTACCGGGAGTGGGGACAGA 6755  
| | | | |  
Db 1890 CGGAGCGAGGGGACTCCGGAGTGTGTGAACCGCTGCTGTACCGGGAGTGGGGACAGA 1831  
| | | | |  
QY 6756 CGGAGCGGACGAGATGGGCTCCAGTCTTCAGGCAACAACAGCCAGCCGACGCTTCT 6815  
| | | | |  
Db 1830 CGGAGCGGACGAGATGGGCTCCAGTCTTCAGGCAACAACAGCCAGCCGACGCTTCT 1771  
| | | | |  
QY 6816 TTAGCAAGCTGACCGAGAGCAACTCCGCAATGGTCAAGTCAAGAAAGCAAGATCAACA 6875  
| | | | |  
Db 1770 TTAGCAAGCTGACCGAGAGCAACTCCGCAATGGTCAAGTCAAGAAAGCAAGATCAACA 1711  
| | | | |  
QY 6876 AGAAGCTGAACAACCAACACCGGAATGAGCTGAATACAAATATACAGCCGCTGGGACGG 6935  
| | | | |  
Db 1710 AGAAGCTGAACAACCAACACCGGAATGAGCTGAATACAAATATACAGCCGCTGGGACGG 1651  
| | | | |  
QY 6936 AGATCTTCAATATGCGCGCATCAACCGGAACAGGCTTATGACTTAATAGAGCAGCGCG 6995  
| | | | |  
Db 1650 AGATCTTCAATATGCGCGCATCAACCGGAACAGGCTTATGACTTAATAGAGCAGCGCG 1591  
| | | | |  
QY 6996 TTAGCAAGACATGCGACACCAACATGGGGCTGAGGCGCATTAATTAGAAAGGCACTCATGG 7055  
| | | | |  
Db 1590 TTAGCAAGACATGCGACACCAACATGGGGCTGAGGCGCATTAATTAGAAAGGCACTCATGG 1531  
| | | | |

QY 7056 GTAATATATGACAGTGGAGAAAGATCCCGCGCTCAGCGCCAAATGCTTTTAACCTCTGA 7115  
| | | | |  
Db 1530 GTAATATATGACAGTGGAGAAAGATCCCGCGCTCAGCGCCAAATGCTTTTAACCTCTGA 1471  
| | | | |  
QY 7116 ATGCCAGTGCAGGCTGCGCGCTGCTATGCTCAATTAACCGTGTCTGACGAGAGTGAAC 7175  
| | | | |  
Db 1470 ATGCCAGTGCAGGCTGCGCGCTGCTATGCTCAATTAACCGTGTCTGACGAGAGTGAAC 1411  
| | | | |  
QY 7176 ACACACTCACTTGGGACAGGTTGGCGGGAGGAGCAAGGCTGTGGACAGCCAGCAGGCC 7235  
| | | | |  
Db 1410 ACACACTCACTTGGGACAGGTTGGCGGGAGGAGCAAGGCTGTGGACAGCCAGCAGGCC 1351  
| | | | |  
QY 7236 GAAAAACCAAGTCCCGGCGCCCGGAGCTGAGACTGAGGAGACCGGACACCTCTGTCTCT 7295  
| | | | |  
Db 1350 GAAAAACCAAGTCCCGGCGCCCGGAGCTGAGACTGAGGAGACCGGACACCTCTGTCTCT 1291  
| | | | |  
QY 7296 CAGTCACTTGGAGGAGACTGCAACCGCGGACGCGCTCAACCAACCGGCTGTGGAGG 7355  
| | | | |  
Db 1290 CAGTCACTTGGAGGAGACTGCAACCGCGGACGCGCTCAACCAACCGGCTGTGGAGG 1231  
| | | | |  
QY 7356 ACAGGCGCTCGTCCGAGGTTCCAGCGCATTCCTCCCAACACCCCGTATCATGGGCTCG 7415  
| | | | |  
Db 1230 ACAGGCGCTCGTCCGAGGTTCCAGCGCATTCCTCCCAACACCCCGTATCATGGGCTCG 1171  
| | | | |  
QY 7416 AGCGGGGTGTCATGAGCTTCCCAACCCGAGCTTCCCGCGGAGCAGCGGCGCCCTCG 7475  
| | | | |  
Db 1170 AGCGGGGTGTCATGAGCTTCCCAACCCGAGCTTCCCGCGGAGCAGCGGCGCCCTCG 1111  
| | | | |  
QY 7476 CTGGCCCCCAACAACGCTTGGAGAGAGAGCCCAAGCCACTGCTCTGCTGCAATCAGAGA 7535  
| | | | |  
Db 1110 CTGGCCCCCAACAACGCTTGGAGAGAGAGCCCAAGCCACTGCTCTGCTGCAATCAGAGA 1051  
| | | | |  
QY 7536 CACTCTCCCAACGAGATCACTCAAGACAGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGG 7595  
| | | | |  
Db 1050 CACTCTCCCAACGAGATCACTCAAGACAGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGG 7501  
| | | | |  
QY 7596 GTTCAGGTCCCAAGAGCCACAGGAAACGCGCTTGCAGAGAGCGGGGCGGCTGCCACTCCC 7655  
| | | | |  
Db 1000 GTTCAGGTCCCAAGAGCCACAGGAAACGCGCTTGCAGAGAGCGGGGCGGCTGCCACTCCC 941  
| | | | |  
QY 7656 CAACCAAGGAAGAGCCCTGAGTCCGCTGCGCTGCATTCATCTGTCCGTCCAGAGCC 7715  
| | | | |  
Db 940 CAACCAAGGAAGAGCCCTGAGTCCGCTGCGCTGCATTCATCTGTCCGTCCAGAGCC 881  
| | | | |  
QY 7716 GGCATCTTGGCTGTCTTAAGCTTAACCTTAACCTTAACTCCCGCGCGCTGCGAG 7775  
| | | | |  
Db 880 GGCATCTTGGCTGTCTTAAGCTTAAAGCTTAACTCCCGCGCGCTGCGAG 821  
| | | | |  
QY 7776 ACCTTAACAGGAGATGTTTAACTGGTGTGCGGAAAGGAGGGGAGCGGAGAGG 7835  
| | | | |  
Db 820 ACCTTAACAGGAGATGTTTAACTGGTGTGCGGAAAGGAGGGGAGCGGAGAGG 761  
| | | | |  
QY 7836 GGCACGCGAGCGGTGTGGAGCCACAACAAGCGGCGAGGGCGGACGAGGACCCAAAGCA 7895  
| | | | |  
Db 760 GGCACGCGAGCGGTGTGGAGCCACAACAAGCGGCGAGGGCGGACGAGGACCCAAAGCA 701  
| | | | |  
QY 7896 GGAATGACACGCACTTCCACGCGCACTGCTTCCCGGAAATGCAATTTGGAACCAAGTCTAA 7955  
| | | | |  
Db 700 GGAATGACACGCACTTCCACGCGCACTGCTTCCCGGAAATGCAATTTGGAACCAAGTCTAA 641  
| | | | |  
QY 7956 ACTGAGCTGCAAGCCCGCGGCTTCCCTCCGCTCCGCTCCCATCCCGCTTAAGCTTGGACA 8015  
| | | | |  
Db 640 ACTGAGCTGCAAGCCCGCGGCTTCCCTCCGCTCCGCTCCCATCCCGCTTAAGCTTGGACA 581  
| | | | |  
QY 8016 GATGAGCGAGGCGCTGTTCAGGCCCCCAAGTGGCTGTTCGAGTCCCAACAATGCCCC 8075  
| | | | |  
Db 580 GATGAGCGAGGCGCTGTTCAGGCCCCCAAGTGGCTGTTCGAGTCCCAACAATGCCCC 521  
| | | | |  
QY 8076 AGCCAAACGAGATTGCTGGAACCAAGTCAAGGCGCAAGTGGCGGAGCAAAAGGCGCAGGTGC 8135  
| | | | |  
Db 520 AGCCAAACGAGATTGCTGGAACCAAGTCAAGGCGCAGGTGGCGGAGCAAAAGGCGCAGGTGC 461  
| | | | |  
QY 8136 GGCCTGGGGGAGAGGATGCTCCGAGGACTGTAAGTGTCTTTTTCACACATCTGTGCGCA 8195  
| | | | |





Dh 901 CGCTATGACCAAGCTCATGGAAGCGCTTGGAAAAAAGGTGGAGCGCATCGAAAACAACCCG 960  
Qy 962 CGCCGGCGGGCCCAAGAGACAAAGGTGCGGAGTACTACGAAAAAGAGTTCCCTGAGATC 1021  
Db 961 CGCCGGCGGGCCCAAGAGACAAAGGTGCGGAGTACTACGAAAAAGAGTTCCCTGAGATC 1020  
Qy 1022 CGCAAGCAGCGCGAGCTGCAAGAGCGCATGCAAGAGAGGTGGGCCAGCGGGCACTGGG 1081  
Db 1021 CCAGAGCAGCGGAGCTGCAAGAGCGCATGCAAGAGAGGTGGGCCAGCGGGCACTGGG 1080  
Qy 1082 CTGTCCATGTCCGCGCCCGCCAGCGACAGAGGTGCAGAGATCATTCGATGCGCTCTCA 1141  
Db 1081 CTGTCCATGTCCGCGCCCGCCAGCGACAGAGGTGCAGAGATCATTCGATGCGCTCTCA 1140  
Qy 1142 GAGCGAGAACCTTGGAGAGAGAGATGCGCGAGCTGGCGGTGATCCCGCCATGCTGTAC 1201  
Db 1141 GAGCGAGAACCTTGGAGAGAGAGATGCGCGAGCTGGCGGTGATCCCGCCATGCTGTAC 1200  
Qy 1202 GACGCTGACCAAGCAGCGCATCAAGTTCAATCAATGACGGGCTTATGCGCCGACCCCATG 1261  
Db 1201 GACGCTGACCAAGCAGCGCATCAAGTTCAATCAATGACGGGCTTATGCGCCGACCCCATG 1260  
Qy 1262 AAGGTGTACAAAGACCGCCAGGTCAATGAACATGTGAGTGTAGCAGAGAGAGACCTTC 1321  
Db 1261 AAGGTGTACAAAGACCGCCAGGTCAATGAACATGTGAGTGTAGCAGAGAGAGACCTTC 1320  
Qy 1322 CGGGAGAAATTGATGAGATCCCAAGAACTTTGGGCTGATGCGATCATTTCCCTGGAGAG 1381  
Db 1321 CGGGAGAAATTGATGAGATCCCAAGAACTTTGGGCTGATGCGATCATTTCCCTGGAGAG 1380  
Qy 1382 AAGACAGTGGCTGAGTGCCTCTCTATTACTACTGACTAAGAGAAATGAGAACTATAAG 1441  
Db 1381 AAGACAGTGGCTGAGTGCCTCTCTATTACTACTGACTAAGAGAAATGAGAACTATAAG 1440  
Qy 1442 AGCTTGTGTAGACGAGACTATGCGCGCGCGCAAGAGCCAGACGACACACGACGACG 1501  
Db 1441 AGCTTGTGTAGACGAGACTATGCGCGCGCGCAAGAGCCAGACGACACACGACGACG 1500  
Qy 1502 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1561  
Db 1501 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
Qy 1562 GATGAGAGAGAGAGAGAGAGAGCGGAGAGAGAGAGAGAGAGCGGAGGTGGAGAAC 1621  
Db 1561 GATGAGAGAGAGAGAGAGAGAGCGGAGAGAGAGAGAGAGAGCGGAGGTGGAGAAC 1620  
Qy 1622 GACAAAGAAAGACTCTCTCAAGAGAGACAGACGACTCAGAGGAGAGACAAAGCAGAG 1681  
Db 1621 GACAAAGAAAGACTCTCTCAAGAGAGAGACAGACGACTCAGAGGAGAGACAAAGCAGAG 1680  
Qy 1682 AAGGAGGCTGTGGCTCTCCAAAGGCGGCAAACTGCGCAACAGCCAGGGAAGAAGCAAGGC 1741  
Db 1681 AAGGAGGCTGTGGCTCTCCAAAGGCGGCAAACTGCGCAACAGCCAGGGAAGAAGCAAGGC 1740  
Qy 1742 CGCATCAACCGGCTCATGTGCTATGAGGCGCAACAGGAGAGGCGCATGCCCGCCAGCAG 1801  
Db 1741 CGCATCAACCGGCTCATGTGCTATGAGGCGCAACAGGAGAGGCGCATGCCCGCCAGCAG 1800  
Qy 1802 AGCGCCGAGCTGGCTCTCAATGAGCTGAATGAAGTTCTCGCTGAGCAGAGAGAGAAATG 1861  
Db 1801 AGCGCCGAGCTGGCTCTCAATGAGCTGAATGAAGTTCTCGCTGAGCAGAGAGAGAAATG 1860  
Qy 1862 GAAACAGCCAAAGAAAGTCTCTTGGAAACGCGCGGCACTGGCTGGCTCCCGCGAGT 1921  
Db 1861 GAAACAGCCAAAGAAAGTCTCTTGGAAACGCGCGGCACTGGCTGGCTCCCGCGAGT 1920  
Qy 1922 GTGGGCTCAAGAGCTGTGCGAGTGTAGAACTTCTTACTTCACTTCAAGAGAGGCGAG 1981  
Db 1921 GTGGGCTCAAGAGCTGTGCGAGTGTAGAACTTCTTACTTCACTTCAAGAGAGGCGAG 1980  
Qy 1982 AACCTTGATGAGATCTTGAGAGACCAAGCTGAAGATGAGAGAGAGAGAGCGCGCG 2041  
Db 1981 AACCTTGATGAGATCTTGAGAGACCAAGCTGAAGATGAGAGAGAGAGAGAGCGCGCG 2040

Qy 2042 AGGAAAGAAAGAAAGCGCGCGCGCGCGCAGCGAGAGGCTGCATTTCCGCGCCGTGTG 2101  
Db 2041 AGGAAAGAAAGAAAGCGCGCGCGCGCGCGCAGCGAGAGGCTGCATTTCCGCGCCGTGTG 2100  
Qy 2102 GAGGATGAGAGAGATGGAAGGCGTGTGGCGTGTGAGCGGAAATGAGAGAGAGATGTGTAGAG 2161  
Db 2101 GAGGATGAGAGAGATGGAAGGCGTGTGGCGTGTGAGCGGAAATGAGAGAGAGATGTGTAGAG 2160  
Qy 2162 GCTGAAGCCTTACATGCGCTCTGAGGAAATGAGGTGCCAGAGGGGAAATGCAAGTGGCCCAAGC 2221  
Db 2161 GCTGAAGCCTTACATGCGCTCTGAGGAAATGAGGTGCCAGAGGGGAAATGCAAGTGGCCCAAGC 2220  
Qy 2222 ACTGTCAACAAAGCTCAGACACCGAGACATCCCTCTCTTCACTGAGGCGGCGCAAG 2281  
Db 2221 ACTGTCAACAAAGCTCAGACACCGAGACATCCCTCTCTTCACTGAGGCGGCGCAAG 2280  
Qy 2282 GACACAGGGCAGAAATGGGCTCAAGGCCCGCAGCCACCTGAGGCGCGAGCGGGCCACCCCA 2341  
Db 2281 GACACAGGGCAGAAATGGGCTCAAGGCCCGCAGCCACCTGAGGCGCGAGCGGGCCACCCCA 2340  
Qy 2342 GGGCCACCCACCCCAACAGAGAGACATCCCGGGCCCCCATTTGAGGCCACCCCGGCTCT 2401  
Db 2341 GGGCCACCCCAACCCCAACAGAGAGACATCCCGGGCCCCCATTTGAGGCCACCCCGGCTCT 2400  
Qy 2402 GAAGCCACCGAGGCCCTTACGCGCCCGCAGCAGCACCCTCATGCGCTCTGCACTCTCTCT 2461  
Db 2401 GAAGCCACCGAGGCCCTTACGCGCCCGCAGCAGCACCCTCATGCGCTCTGCACTCTCTCT 2460  
Qy 2462 GTGTGTCCCAAGAGAGAGAGAGAGAGAGACCGCAGCAGCGGCCCAAGTGAAGAGAGG 2521  
Db 2461 GTGTGTCCCAAGAGAGAGAGAGAGAGAGACCGCAGCAGCGGCCCAAGTGAAGAGAGG 2520  
Qy 2522 GAGAGAGAGAAAGCCCCCGCGCTGAGAGAGCTGTGAGCTGTGACACAGAGAGAGGCGGAGAG 2581  
Db 2521 GAGAGAGAGAAAGCCCCCGCGCTGAGAGAGCTGTGAGCTGTGACACAGAGAGAGGCGGAGAG 2580  
Qy 2582 CCGGTCAAGAGAGAGTGTGACGAGAGAGCCGAGAGGGGCGGGCCAAAGGGCAAGAGACGCG 2641  
Db 2581 CCGGTCAAGAGAGAGTGTGACGAGAGAGCCGAGAGGGGCGGGCCAAAGGGCAAGAGACGCG 2640  
Qy 2642 GAGGCGCTGAGGGCCACGCGCCGAGGGGGCGCTCAAGGCGAGAGAGAGAGGGCGGGAGC 2701  
Db 2641 GAGGCGCTGAGGGCCACGCGCCGAGGGGGCGCTCAAGGCGAGAGAGAGAGGGCGGGAGC 2700  
Qy 2702 GGCAGGGCCACACTGTGCAAGAGCTGTGGGCGCCCCCAGAGACGCGACTTCAAGTGTAC 2761  
Db 2701 GGCAGGGCCACACTGTGCAAGAGCTGTGGGCGCCCCCAGAGACGCGACTTCAAGTGTAC 2760  
Qy 2762 TGCATGTGAGAGAGAGTGTGATGAGGCGGAGGGCGGCGAGCAAGAAACGGGCTGTGCCCA 2821  
Db 2761 TGCATGTGAGAGAGAGTGTGATGAGGCGGAGGGCGGCGAGCAAGAAACGGGCTGTGCCCA 2820  
Qy 2822 AAGGCCAGGCTCTCAACCCCGACTGGCGACCCCGGGGCAATGCTCAACCCAGAGAGCA 2881  
Db 2821 AAGGCCAGGCTCTCAACCCCGACTGGCGACCCCGGGGCAATGCTCAACCCAGAGAGCA 2880  
Qy 2882 CTGAGACTGAGCAGCTGAAGCAGCGAGCGGCTGTGCCATCTCCCTCATCAAGGTCAACAAA 2941  
Db 2881 CTGAGACTGAGCAGCTGAAGCAGCGAGCGGCTGTGCCATCTCCCTCATCAAGGTCAACAAA 2940  
Qy 2942 GTTCATAGGCGCGCGCGGAGAGACGAGCTGCCAACCAAGCAGGCTGCCCGACCCCAAGC 3001  
Db 2941 GTTCATAGGCGCGCGCGGAGAGACGAGCTGCCAACCAAGCAGGCTGCCCGACCCCAAGC 3000  
Qy 3002 CCAACGCAAAACCTTGAAGCGGAGAGAGCAGCGCCCTCAGCAGCCTTGTGACAGCGCCCGG 3061  
Db 3001 CCAACGCAAAACCTTGAAGCGGAGAGAGCAGCGCCCTCAGCAGCCTTGTGAGAGCGCCCGG 3060  
Qy 3062 GGCAGAGCAGAGAGCGCCGAGACCCCGCGCGACAAAGAGGCTTTGCGAGCGGAGGCGCGAG 3121  
Db 3061 GGCAGAGCAGAGAGCGCCGAGACCCCGCGCGACAAAGAGGCTTTGCGAGCGGAGGCGCGAG 3120

QY	3122	AAGCTGCTGGGGGACCCCCCTTGGCTGGACCTTCGGGCTGGCCCTTCCCGTGGCCCCCGCT	3128
Dp	3121	AAGCTGCTGGGGGACCCCCCTTGGCTGGACCTTCGGGCTGGCCCTTCCCGTGGCCCCCGCT	3180
QY	3182	GAGGTGATCAAGAGCTCCCGCATGACCCCGGAACCCCTCAGGCTTCTCTTACGCTCACT	3241
Dp	3181	GAGGTGATCAAGAGCTCCCGCATGACCCCGGAACCCCTCAGGCTTCTCTTACGCTCACT	3240
QY	3242	GGTACCCCATGCCCCCTGGGCTTCATGACATGCGCGGCCCCGTCTGCGGCCCCACCC	3301
Dp	3241	GGTACCCCATGCCCCCTGGGCTTCATGACATGCGCGGCCCCGTCTGCGGCCCCACCC	3300
QY	3302	ACCATCTCCAAACCCGCTCCCTCCATCTCCCTGCGAAGACACCCAGGCTCTGAGAGG	3361
Dp	3301	ACCATCTCCAAACCCGCTCCCTCCATCTCCCTGCGAAGACACCCAGGCTCTGAGAGG	3360
QY	3362	CAATAGGTGCTATCTCCCAAGAAATGCGGTCCAGCTCCAGTCCGCTACTCAGAGAT	3421
Dp	3361	CAATAGGTGCTATCTCCCAAGAAATGCGGTCCAGCTCCAGTCCGCTACTCAGAGAT	3420
QY	3422	GCCAAAGCCCCGTGGGCGCTGTCAACATGGGGCTGCCCTGCCCCATGAAACCCAAAAAG	3481
Dp	3421	GCCAAAGCCCCGTGGGCGCTGTCAACATGGGGCTGCCCTGCCCCATGAAACCCAAAAAG	3480
QY	3482	CTGGCACCCCTTCACTCGGATGTAAGAGAGACGACGTCTCCCAACGGGGGCAAGCTGGGCA	3541
Dp	3481	CTGGCACCCCTTCACTCGGATGTAAGAGAGACGACGTCTCCCAACGGGGGCAAGCTGGGCA	3540
QY	3542	CCGGAGAGCTGGGGGTGCCCAACAGCCGAGAGGGCGTCCGTGTGAGAGGACAAGCTTG	3601
Dp	3541	CCGGAGAGCTGGGGGTGCCCAACAGCCGAGAGGGCGTCCGTGTGAGAGGACAAGCTTG	3600
QY	3602	GGCTCAGTTCGGGGCGAAGACATCAACAAAGCATTTCCACAGACAACGGGTGCCCTTGGAC	3661
Dp	3601	GGCTCAGTTCGGGGCGAAGACATCAACAAAGCATTTCCACAGACAACGGGTGCCCTTGGAC	3660
QY	3662	AGCGGCATCAATACCGGCGCTCCATCAACCAACGGACGCCAGCTGAGTCTTGTACAAAG	3721
Dp	3661	AGCGGCATCAATACCGGCGCTCCATCAACCAACGGACGCCAGCTGAGTCTTGTACAAAG	3720
QY	3722	GGCACCATCACACAGGATCATCGGCGAGAGACAGCCGAGTCCGTTGGAACGGGCGGGAG	3781
Dp	3721	GGCACCATCACACAGGATCATCGGCGAGAGACAGCCGAGTCCGTTGGAACGGGCGGGAG	3780
QY	3782	GACAGCCTGCCCAAGGGCCACGTCACTTACAGAAAGGCAAGAGGGCAAGTCTTGCTAT	3841
Dp	3781	GACAGCCTGCCCAAGGGCCACGTCACTTACAGAAAGGCAAGAGGGCAAGTCTTGCTAT	3840
QY	3842	GAGGTGGCATGTCTGTGACCAATGCTTCAGAGAGACGGCAGAGACGCTCAGAGACC	3901
Dp	3841	GAGGTGGCATGTCTGTGACCAATGCTTCAGAGAGACGGCAGAGACGCTCAGAGACC	3900
QY	3902	CCCATGAGAGGGCGCCCAAGGGCACATGACATGATGGAGGGCGGGTGGGACGA	3961
Dp	3901	CCCATGAGAGGGCGCCCAAGGGCACATGACATGATGAGAGGGCGGGTGGGACGA	3960
QY	3962	GCCATCTCTCAGCCAGCATGAAAGTCTCATGAGGCGGTGCATCTCCGCCGAGGACAC	4021
Dp	3961	GCCATCTCTCAGCCAGCATGAAAGTCTCATGAGGCGGTGCATCTCCGCCGAGGACAC	4020
QY	4022	AGCCCCACCACTTCAAGAGACGACACATCCGCGGGTCCATCAACAAGGATCCCT	4081
Dp	4021	AGCCCCACCACTTCAAGAGACGACACATCCGCGGGTCCATCAACAAGGATCCCT	4080
QY	4082	CGGTCCTACGGGAGGCAACGAGACATACCTGGCGTCCGAGAGGCCAAGCTCTTAAAGGG	4141
Dp	4081	CGGTCCTACGGGAGGCAACGAGACATACCTGGCGTCCGAGAGGCCAAGCTCTTAAAGGG	4140
QY	4142	GAGGGCAAGCTTCGGCCCCCAACGGCTTCAAGGAGCTGACCGAGGCTTACAAGAGCGAG	4201
Dp	4141	GAGGGCAAGCTTCGGCCCCCAACGGCTTCAAGGAGCTGACCGAGGCTTACAAGAGCGAG	4200
QY	4202	GCCCTGGGCCCCCTTGAAGTGAAGCCGGGCCCATGAGGGCTTGTTGGCCACGTTGAAGAG	4261

Db	4201	GCCCTGGGGCCCCCTGAAAGCTGAAGCCGGGCCCATGAGGGCTGGTGGCCACGGTGAAGAG	4260
OY	4262	GGGGGCGGCTTCATTCATAGAGATCCCGCGAGAGCTGCGCACGCTCGAGCTGGCC	4321
Db	4261	GGGGGCGGCTTCATTCATAGAGATCCCGCGAGAGACTCGGGGACACGCCCCAGCTGGCC	4320
OY	4322	CTGGCCCCGGGGCGGCTCAAGAGAGGGGCTCCATACGACAGAGGACCCCCGTCMAAGTACGAC	4381
Db	4321	CTGGCCCCGGGGCGGCTCAAGAGAGGGGCTCCATACGACAGAGGACCCCCGTCMAAGTACGAC	4380
OY	4382	ACCGGCGGCTTCACACATGSGCTCCAAAAAGACGAGTACGCTCCCTCATCGGAGCCCC	4441
Db	4381	ACCGGCGGCTTCACACATGSGCTCCAAAAAGACGAGTACGCTCCCTCATCGGAGCCCC	4440
OY	4442	GGCGGAGCGTTTCCACCCCGTGCACCCGCTGATGTATGAGCGGACCGCGGAGCATGGAA	4501
Db	4441	GGCGGAGCGTTTCCACCCCGTGCACCCGCTGATGTATGAGCGGACCGCGGAGCATGGAA	4500
OY	4502	CGTGCCTGCTACGAGAGAGAGCTGAAGAGGCGGCGACGAGACCGGACAGCACTCGGGGGG	4561
Db	4501	CGTGCCTGCTACGAGAGAGAGCTGAAGAGGCGGCGACGAGACCGGACAGCACTCGGGGGG	4560
OY	4562	TTCAATTGGCGGGCGGGCGGCGGCTCATTTTGGCTGAGCTGSGGTAAAGCGGGGAGAGCCCC	4621
Db	4561	TTCAATTGGCGGGCGGGCGGCGGCTCATTTTGGCTGAGCTGSGGTAAAGCGGGGAGAGCCCC	4620
OY	4622	CTGACCTATGAGAGACACGAGGGGACCCCTTTGGCCGGCACCTCCACAGAGTTTGGCCGTG	4681
Db	4621	CTGACCTATGAGAGACACGAGGGGACCCCTTTGGCCGGCACCTCCACAGAGTTTGGCCGTG	4680
OY	4682	ACCATGGCGGAGGCCACGCGCGCGCTGACAGAGAGGAGAGCTTTCTGTACAGCAAGGATCC	4741
Db	4681	ACCATGGCGGAGGCCACGCGCGCGCTGACAGAGAGGAGAGCTTTCTGTACAGCAAGGATCC	4740
OY	4742	CAGGACCGGAAGGCTGAGTGAAGGCTCGTGAATGCCAATGCCGCCGACACGACCCGTG	4801
Db	4741	CAGGACCGGAAGGCTGAGTGAAGGCTCGTGAATGCCAATGCCGCCGACACGACCCGTG	4800
OY	4802	CCCGAGACCAACCCACACCCCATCTCGCCCTATGAGACCTGTCTTGAGGAGCGTGAATGAC	4861
Db	4801	CCCGAGACCAACCCACACCCCATCTCGCCCTATGAGACCTGTCTTGAGGAGCGTGAATGAC	4860
OY	4862	GTTGACCTGTATGCGAGCCACATCCCTCTGGGCTTGCACCCACTTCATACCCCGCGGC	4921
Db	4861	GTTGACCTGTATGCGAGCCACATCCCTCTGGGCTTGCACCCACTTCATACCCCGCGGC	4920
OY	4922	ATCCCTTTGAGACCGAGCGGCTGTACTACCTGCGCGGACACTGTGGCCGCCAACCCGACC	4981
Db	4921	ATCCCTTTGAGACCGAGCGGCTGTACTACCTGCGCGGACACTGTGGCCGCCAACCCGACC	4980
OY	4982	TACCGGACCTGTATACCAACCTTACTCATCCGCGGCTTACCCGACACGCGCGGCGCTGAG	5041
Db	4981	TACCGGACCTGTATACCAACCTTACTCATCCGCGGCTTACCCGACACGCGCGGCGCTGAG	5040
OY	5042	AACCGGAGACCATCATCAATGAATGATCATCACTGTGAGCAAGTGCACCAACACGCGC	5101
Db	5041	AACCGGAGACCATCATCAATGAATGATCATCACTGTGAGCAAGTGCACCAACACGCGC	5100
OY	5102	ACCGGCAATGGCCCAAGAGAGTGAATGTGTGAGGGGGCTTCCGCCCGGAGAGTCCCTGCGT	5161
Db	5101	ACCGGCAATGGCCCAAGAGAGTGAATGTGTGAGGGGGCTTCCGCCCGGAGAGTCCCTGCGT	5160
OY	5162	GCACTGAACACTAGCTGGGGGTCCCGAGAGGACATATGACCTGTGCCAATGTGCACACTTG	5221
Db	5161	GCACTGAACACTAGCTGGGGGTCCCGAGAGGACATATGACCTGTGCCAATGTGCACACTTG	5220
OY	5222	CTGTGTCTGTGCCCCCGACACCAAGGACCCCGACGCCACCGGCATGAGACCGCTTGTGCTAC	5281
Db	5221	CTGTGTCTGTGCCCCCGACACCAAGGACCCCGACGCCACCGGCATGAGACCGCTTGTGCTAC	5280
OY	5282	CTCCCAACCGGGCCCCAGCCCTTACAGAGCGGCGCACAGAGCTTCCCACTTTCGCCAGGA	5341

Db 5281 CTCGCCGCGCCGACGCCCTTCAGACGCGCCAGACGACTCCCACTTCCCGCAGGA 5340  
Qy 5342 GGTCAACAACATTGACAAACCAACCAACCAACGTCCTGTCGAGCGGAGCGAGACCG 5401  
Db 5341 GGTTCACAACAATTGACAAACCAACCAACCAACGTCCTGTCGAGCGGAGCGAGACCG 5400  
Qy 5402 GATCGAGACCGGAGACCGGAGTCGGAGCGGAGAAAGTCCATCTCTCACTCTCAACGACG 5461  
Db 5401 GATCGAGACCGGAGACCGGAGTCGGAGCGGAGAAAGTCCATCTCTCACTCTCAACGACG 5460  
Qy 5462 GTGAGACACGACCCATCTGAGAACCTGTGTACAGACAGACAGCGCGACAGCGGACG 5521  
Db 5461 GTGAGACACGACCCATCTGAGAACCTGTGTACAGACAGACAGCGCGACAGCGGACG 5520  
Qy 5522 AGCGGCGGAGGTGAGGAGGAGACAGACGCGCGCGCTCCCACTCCATGCGCCACAGAC 5581  
Db 5521 AGCGGCGGAGGTGAGGAGGAGACAGACGCGCGCGCTCCCACTCCATGCGCCACAGAC 5580  
Qy 5582 TCGCCCATCTCCCTCGGACCCAGATGCGCTTCAGACAGACCCAGTGTCTTCAACA 5641  
Db 5581 TCGCCCATCTCCCTCGGACCCAGATGCGCTTCAGACAGACCCAGTGTGTCAACA 5640  
Qy 5642 ACAGGATGAGAGGTATCATCACCGCTGTGAGCCAGACAGACCGCTCTGAGGTCC 5701  
Db 5641 ACAGGATGAGAGGTATCATCACCGCTGTGAGCCAGACAGACCGCTCTGAGGTCC 5700  
Qy 5702 ACCTCACCTCTCAACCGGTGCGCCAGGTCACATTCGCACTGCGCACCCATGCGCA 5761  
Db 5701 ACCTCACCTCTCAACCGGTGCGCCAGGTCACATTCGCACTGCGCACCCATGCGCA 5760  
Qy 5762 CTGAGCGGACCCCTCGATGAGGTCTACCCCTACCTCATGAGGCGCTCTTGTGCTCCCAAG 5821  
Db 5761 CTGAGCGGACCCCTCGATGAGGTCTACCCCTACCTCATGAGGCGCTCTTGTGCTCCCAAG 5820  
Qy 5822 GAGGCGCGCGGAGTGCCTCGGACAGCGGCGCCCGACAGACACCGGCAATGCTTCTTC 5881  
Db 5821 GAGGCGCGCGGAGTGCCTCGGACAGCGGCGCCCGACAGACACCGGCAATGCTTCTTC 5880  
Qy 5882 GACAGGCGCGGCGGCTCGGAGCGGAGCGCGCTCTCCCGCCAGAAAGGCTCGGAG 5941  
Db 5881 GACAGGCGCGGCGGCTCGGAGCGGAGCGCGCTCTCCCGCCAGAAAGGCTCGGAG 5940  
Qy 5942 CCCCGGCGCGTATGCTCTCTCTCTGAGCACGACCAACATGACCGGACCCCTGCGAAG 6001  
Db 5941 CCCCGGCGCGTATGCTCTCTCTCTGAGCACGACCAACATGACCGGACCCCTGCGAAG 6000  
Qy 6002 AACCTGCACTTCACGACGCGACCGCGGACCGGCGCGGCACTGCGCTCGCTCGAC 6061  
Db 6001 AACCTGCACTTCACGACGCGACCGCGGACCGGCGCGGCACTGCGCTCGCTCGAC 6060  
Qy 6062 CGGACCGCGGAAAGACCTAAAGTAAACCTTTTCAATCAAGAACTGGAATCTCGTTC 6121  
Db 6061 CGGACCGCGGAAAGACCTAAAGTAAACCTTTTCAATCAAGAACTGGAATCTCGTTC 6120  
Qy 6122 CTGAGTTACACGCGACGACTACAGCCCGGAAAGGAGTGAAGCGCTGACCGCTGAGAC 6181  
Db 6121 CTGAGTTACACGCGACGACTACAGCCCGGAAAGGAGTGAAGCGCTGAGCGCTGAGAC 6180  
Qy 6182 TCACCCAGTCTGACCGACGACAAAGGAGTCCCAAGCACCTGAGAAAGTTCGACAAAGAC 6241  
Db 6181 TCACCCAGTCTGACCGACGACAAAGGAGTCCCAAGCACCTGAGAAAGTTCGACAAAGAC 6240  
Qy 6242 CACCTGAGAGGAGAGTGTGCGCCCAAGGACGACGCGCGGTGAAGCTTGTGCGGAGGCGC 6301  
Db 6241 CACCTGAGAGGAGAGTGTGCGCCCAAGGACGACGCGCGGTGAAGCTTGTGCGGAGGCGC 6300  
Qy 6302 GCGCACTTCCACACCTGCGCGCTGCTGAGAGCGACGCTCTGTCAAGCGCGCTGCTC 6361  
Db 6301 GCGCACTTCCACACCTGCGCGCTGCTGAGAGCGACGCTCTGTCAAGCGCGCTGCTC 6360  
Qy 6362 CAGACCGCGCGGAGGTCAAAGGTACACAGCGGAGTGTCACTTGTGCGCCAGACATCACT 6421  
Db 6361 CAGACCGCGCGGAGGTCAAAGGTACACAGCGGAGTGTCACTTGTGCGCCAGACATCACT 6420

Qy 6422 GAGTCAATCACAGGACTTACACCCGCGACCAACCCAGAGCTCAGCGCACCCCTGCGC 6481  
Db 6421 GAGTCAATCACAGGACTTACACCCGCGACCAACCCAGAGCTCAGCGCACCCCTGCGC 6480  
Qy 6482 GCGCCCTCTACTCTTCCCTGAGGAGCACTGCGCTCTGGAACCTCGCGCGCCACCC 6541  
Db 6481 GCGCCCTCTACTCTTCCCTGAGGAGCACTGCGCTCTGGAACCTCGCGCGCCACCC 6540  
Qy 6542 AGTGAACCTTACCTCCCGCGCCCGGACCAATGAGCGCGCGCGCTGAGCTCCCCCAAGC 6601  
Db 6541 AGTGAACCTTACCTCCCGCGCCCGGACCAATGAGCGCGCGCGCTGAGCTCCCCCAAGC 6600  
Qy 6602 GAAAGGAGCAAGAGTCTCAAGAGCCAAACAAAGCTCGGCTCTGAGGTGTGAGAC 6661  
Db 6601 GAAAGGAGCAAGAGTCTCAAGAGCCAAACAAAGCTCGGCTCTGAGGTGTGAGAC 6660  
Qy 6662 GGTATTGAACCTGTGTCCCAACCGAGGAGCATGACGAGCCAGAGGACCTCCGAGTGCCT 6721  
Db 6661 GGTATTGAACCTGTGTCCCAACCGAGGAGCATGACGAGCCAGAGGACCTCCGAGTGCCT 6720  
Qy 6722 GTGTACCGCGCTGTGTACCGGAGTGGGAGAACAGCGGAGCCCAAGACGATGGCTTCAAG 6781  
Db 6721 GTGTACCGCGCTGTGTACCGGAGTGGGAGAACAGCGGAGCCCAAGACGATGGCTTCAAG 6780  
Qy 6782 TCTCCAGGCAACACGACGCGCGACGCTTCTTCAAGACGTGACCGAGACACTCC 6841  
Db 6781 TCTCCAGGCAACACGACGCGCGACGCTTCTTCAAGACGTGACCGAGACACTCC 6840  
Qy 6842 GCCATGTCAGTCCAGAGGACGAGATCAACAAAGCTGAACCCACACCGGAAT 6901  
Db 6841 GCCATGTCAGTCCAGAGGACGAGATCAACAAAGCTGAACCCACACCGGAAT 6900  
Qy 6902 GAGCTGAATTAATATTCAGCCAGCTGGGAGCGGAGATTTCAATATGCGCGCATACAC 6961  
Db 6901 GAGCTGAATTAATATTCAGCCAGCTGGGAGCGGAGATTTCAATATGCGCGCATACAC 6960  
Qy 6962 GGAACGAGCTTATGACCTATGAAGCCAGCGGAGTGAAGAACTGCAAGCAACATG 7021  
Db 6961 GGAACGAGCTTATGACCTATGAAGCCAGCGGAGTGAAGAACTGCAAGCAACATG 7020  
Qy 7022 GGGCTGAGAGGCAATTAATGAAGGCACTCATGGTAAATATGACAGTGGAGAGTCC 7081  
Db 7021 GGGCTGAGAGGCAATTAATGAAGGCACTCATGGTAAATATGACAGTGGAGAGTCC 7080  
Qy 7082 CGCGCGCTCAGCGGCAATGCTTTTAAACCTTGAATGCAAGTGCACGCTGCGCTGCT 7141  
Db 7081 CGCGCGCTCAGCGGCAATGCTTTTAAACCTTGAATGCAAGTGCACGCTGCGCTGCT 7140  
Qy 7142 ATGCCCATTAACGCGTGTGAGCGGAGTGAACAACACTCAACCTGCGCAGGTGCGGCG 7201  
Db 7141 ATGCCCATTAACGCGTGTGAGCGGAGTGAACAACACTCAACCTGCGCAGGTGCGGCG 7200  
Qy 7202 GGGAGAGCCAGAGTCTTGTGAGACCCAGACCGGAGGAGGAGGAGTCCCGGCGCGGCG 7261  
Db 7201 GGGAGAGCCAGAGTCTTGTGAGACCCAGACCGGAGGAGGAGGAGTCCCGGCGCGGCG 7260  
Qy 7262 CTGGCATTTGGGAGACCGGCACTCTGTGTCTCTCACTGATGCACTCGAGGAGGAGCTGCAAC 7321  
Db 7261 CTGGCATTTGGGAGACCGGCACTCTGTGTCTCTCACTGATGCACTCGAGGAGGAGCTGCAAC 7320  
Qy 7322 GCGCGGAGCGGCTCAACCAACGCGGTGTGGAGGAGCAAGGCGCTCGCGCAGGTTCCAG 7381  
Db 7321 GCGCGGAGCGGCTCAACCAACGCGGTGTGGAGGAGCAAGGCGCTCGCGCAGGTTCCAG 7380  
Qy 7382 CCATTCCTCTCAACCCCTCTGATCATGCGGCTGACAGCGGAGTGTACATGAGCTTCCCAACC 7441  
Db 7381 CCATTCCTCTCAACCCCTCTGATCATGCGGCTGACAGCGGAGTGTACATGAGCTTCCCAACC 7440  
Qy 7442 CCAACCGGCTCTCCCGCGGAGCGGCGCTCTGCGTGGCGCCCAACAGCTCTGGAGACGAG 7501  
Db 7441 CCAACCGGCTCTCCCGCGGAGCGGCGCTCTGCGTGGCGCCCAACAGCTCTGGAGACGAG 7500

QY 7502 GAGCCCAAGCCACTGCTCTGCGAGTACGAGACACTCTCCGACAGCGAGTGA 7555  
DB 7501 GAGCCCAAGCCACTGCTCTGCGAGTACGAGACACTCTCCGACAGCGAGTGA 7554

RESULT 14

ID ADL13812 standard; DNA; 7524 BP.

AC ADL13812;

DT 06-MAY-2004 (first entry)

DE Osteoarthritis-associated polymorphic nucleotide #344.

KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;

KW joint space narrowing; osteophyte development; joint pain;

KW osteoarthritis; SNP; single nucleotide polymorphism.

OS Homo sapiens.

PN WO2003054166-A2.

PD 03-JUL-2003.

PE 19-DEC-2002; 2002WO-US041225.

PR 20-DEC-2001; 2001US-0342603P.

PA (INCY-) INCYTE GENOMICS INC.

P1 Jones KA, Schafer A;

DR WPI; 2003-559141/52.

PT Determining susceptibility of an individual to joint space narrowing, osteophyte development and/or joint pain comprises identifying whether the individual has at least one polymorphism in a polynucleotide encoding a protein.

PS Disclosure: SEQ ID NO 344; 297bp; English.

CC The invention relates to a method of determining susceptibility of an individual to joint space narrowing and/or osteophyte development and/or joint pain comprising identifying whether the individual has at least one polymorphism in a polynucleotide encoding at least one of the protein listed in the specification. The methods, composition and agent are useful for modulating the susceptibility of an individual to joint space narrowing and/or osteophyte development and/or joint pain that is associated with a disease, preferably osteoarthritis. The cell line and the non-human animal are useful for screening for an agent for diagnosing an individual having susceptibility to joint space narrowing and/or osteophyte development and/or joint pain. This sequence corresponds to the polynucleotide encoding a protein listed in the specification. (Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences).

XX Sequence 7524 BP; 1636 A; 2728 C; 2213 G; 947 T; 0 U; 0 Other;

QY Query Match 85.7%; Score 7337.6; DB 10; Length 7524;

DB Best Local Similarity 98.5%; Pred. No. 0; Matches 7466; Conservative 0; Mismatches 34; Indels 78; Gaps 3;

QY 2 ATGTGGGCTCCACAGCTTGTGGACAGAGTGAAGGCGACGAGCCCGCTACCG 61

DB 1 ATGTGGGCTCCACAGCTTGTGGACAGAGTGAAGGCGACGAGCCCGCTACCG 60

QY 62 CCCACAGACCTTTCTTACCCAGTGAAGATGCGCCGAGACGACGAGAGCTGGCTCTG 121

DB 61 CCCACAGACCTTTCTTACCCAGTGAAGATGCGCCGAGACGACGAGAGCTGGCTCTG 120

QY 122 GAGTACGACGACCACTCCCGCGACTATAGCTTCCACTGTGCGCGGCTCATATCCAG 181

DB 121 GAGTACGACGACCACTCCCGCGACTATAGCTTCCACTGTGCGCGGCTCATATCCAG 180

QY 182 CCCACAGACGAGGCTCTCCCTGCTGTAGATTCCAGCCCGGAGTAAAGGCTCCAG 241

DB 181 CCCACAGACGAGGCTCTCCCTGCTGTAGATTCCAGCCCGGAGTAAAGGCTCCAG 240

QY 242 GAGTCCACTGTCGCGCCAGAGTCCCACTCATCTGCTCCGAGCTGGGAGTCCAGATG 301

DB 241 GAGTCCACTGTCGCGCCAGAGTCCCACTCATCTGCTCCGAGCTGGGAGTCCAGATG 300

QY 302 GAGTTCAATTGAAAGAGAGCGCTCTGCTAGAGCTGCTGAGCCCTCTGCTGGAGCG 361

DB 301 GAGTTCAATTGAAAGAGAGCGCTCTGCTAGAGCTGCTGAGCCCTCTGCTGGAGCG 360

QY 362 TCACCCCTGCTGCGCCAGAGCTCCGAGATCTGAAGACCTCAACCAAGACCGTAGC 421

DB 361 TCACCCCTGCTGCGCCAGAGCTCCGAGATCTGAAGACCTCAACCAAGACCGTAGC 420

QY 422 CTGACGAGGCAAGCTTGAACCGGTGTCTCCGCCAGCCCGCCGCACTGACCTTGAGCTG 481

DB 421 CTGACGAGGCAAGCTTGAACCGGTGTCTCCGCCAGCCCGCCGCACTGACCTTGAGCTG 480

QY 482 GAGCTGTGCTCCGACCGGCTGTCCAAAGAGAGAGCTGATCCAGAAACATGAGCCGTGAC 541

DB 481 GAGCTGTGCTCCGACCGGCTGTCCAAAGAGAGAGCTGATCCAGAAACATGAGCCGTGAC 540

QY 542 CGAGAGATCAACATGATGAGACAGAGATCTCTAAGCTGAAGAAAGAACAGACAGCTG 601

DB 541 CGAGAGATCAACATGATGAGACAGAGATCTCTAAGCTGAAGAAAGAACAGACAGCTG 600

QY 602 GAGGAGAGGCTGTGCCAAGCCGCGAGCTGAGAAAGCCGTGTCAACGCGCCCATTCAG 661

DB 601 GAGGAGAGGCTGTGCCAAGCCGCGAGCTGAGAAAGCCGTGTCAACGCGCCCATTCAG 660

QY 662 TCGAAGCACCCGACGCTGTGTCAATCATCTTACACAGGAACCGGAAGAAGCTGAAAGCT 721

DB 661 TCGAAGCACCCGACGCTGTGTCAATCATCTTACACAGGAACCGGAAGAAGCTGAAAGCT 720

QY 722 GCACATCGGATTTCTGGAAGGCTGTGGGCGCCGAGGTGAGTGCCTGTACAAACGAGCC 781

DB 721 GCACATCGGATTTCTGGAAGGCTGTGGGCGCCGAGGTGAGTGCCTGTACAAACGAGCC 780

QY 782 TCCGACACCCGCGAGATATCATGAGAAACATCAAAATTAACAGGCGATCGGAAGAGCTA 841

DB 781 TCCGACACCCGCGAGATATCATGAGAAACATCAAAATTAACAGGCGATCGGAAGAGCTA 840

QY 842 ATCTTGTACTTCAAGAGAGAGATCAAGCTGTGGAACCAATGGAACAGAAATTCTGCCAG 901

DB 841 ATCTTGTACTTCAAGAGAGAGATCAAGCTGTGGAACCAATGGAACAGAAATTCTGCCAG 900

QY 902 CGCTATGACAGCTCATGAGAGGCTTGGAAAAAAGGTGAGCGCATGAAAAACAACCG 961

DB 901 CGCTATGACAGCTCATGAGAGGCTTGGAAAAAAGGTGAGCGCATGAAAAACAACCG 960

QY 962 CGCGGCGGCGCCAAAGAGAGCAAGGTGCGAGTACTACGAAAAAGCAAGTCTCTGAGATC 1021

DB 961 CGCGGCGGCGCCAAAGAGAGCAAGGTGCGAGTACTACGAAAAAGCAAGTCTCTGAGATC 1020

QY 1022 CGCAAGACGCGGAGCTGCAAGAGCGCATGCAAGACAGAGGTGGGCTCAAGCGGCGAGTGGG 1081

DB 1021 CGCAAGACGCGGAGCTGCAAGAGCGCATGCAAGAGCGCATGCAAGAGGTGGGCTCAAGCGGCGAGTGGG 1077

QY 1082 CTGTCCATGTGGCGCGCGCCGAGCGAGACAGAGGTGTCAGAGATCATGATGAGCTCTCA 1141

DB 1078 CTGTCCATGTGGCGCGCGCCGAGCGAGACAGAGGTGTCAGAGATCATGATGAGCTCTCA 1137

QY 1142 GAGCAGAGAACTTGAAGAACAGATGCGCAAGCTGCGCGTGAATCCGCGCATGTGTGAC 1201

DB 1138 GAGCAGAGAACTTGAAGAACAGATGCGCAAGCTGCGCGTGAATCCGCGCATGTGTGAC 1197

QY 1202 GACGTTGACGAGCGGATCAAGTTTATCAATGAACGGGCTTAAGCCGACCCCATG 1261





QY 3398 CTCAAGTTCGGTACTCAGAGCATCCAAAGSCCCCGTGGGCTCTGTCAACATG9GGGCTG 3457  
Db 3367 CTCCAGTCCCGTACTCAGAGCATCCAAAGSCCCCGTGGGCTCTGTCAACATG9GGGCTG 3426  
QY 3458 CCCCGGCTATG9AGCCCAAAAAGTGGCACTTCAAGGAGTGAAGCAGAGAGAGCTG 3517  
Db 3427 CCGCTGGCCATG9AGCCCAAAAAGTGGCACTTCAAGGAGTGAAGCAGAGAGAGCTG 3486  
QY 3518 TCCCCACGGGGCAGAGCTGGGGCAACCGAGAGCTGGGGGTGTCACAGCCCAAGAGGCG 3577  
Db 3487 TCCCCACGGGGCAGAGCTGGGGCAACCGAGAGCTGGGGGTGTCACAGCCCAAGAGGCG 3546  
QY 3578 TCCGTCGTGAAGAGG9ACAGCTTGGGCTCAATTCCGGGCGAAGAGATCAACCAAGGCAAT 3637  
Db 3547 TCCGTCGTGAAGAGG9ACAGCTTGGGCTCAATTCCGGGCGAAGAGATCAACCAAGGCAAT 3606  
QY 3658 CCCAGCACAAGGGGTGCTCTGGGACAGCGCCATCATATCCGGGGCTCCATCAACCAAGGCG 3697  
Db 3607 CCCAGCACAAGGGGTGCTCTGGGACAGCGCCATCATATCCGGGGCTCCATCAACCAAGGCG 3666  
QY 3698 ACGCAGCTGAAGTCTGTACAAAGG9CACATCAACAGAGATCATCGAGCAGAGACAGCCCG 3757  
Db 3667 ACGCAGCTGAAGTCTGTACAAAGG9CACATCAACAGAGATCATCGAGGAGACAGCCCG 3726  
QY 3758 AGTCGCTTGAAGCCGGGCGGGAGAGACAGCTTGCCCAAGG9CACATCTTAAGAGGCG 3817  
Db 3727 AGTCGCTTGAAGCCGGGCGGGAGAGACAGCTTGCCCAAGG9CACATCTTAAGAGGCG 3786  
QY 3818 AAGAAAGG9CACAGTGTGCTCATGAGGGGTGGCATGTCTGTGACCCAGTGTCCCAAGGAG 3877  
Db 3787 AAGAAAGG9CACAGTGTGCTCATGAGGGGTGGCATGTCTGTGACCCAGTGTCCCAAGGAG 3846  
QY 3878 GACG9CAGAAAGAGCTCAGAGACCCGCCCATGAGACG9CGGCCCCCAAGG9CACATTAAGC 3937  
Db 3847 GACG9CAGAAAGAGAGCTCAGAGACCCGCCCATGAGAGAGGCGGCCCCCAAGG9CACATTAAGC 3906  
QY 3938 ATGATGAGAGG9CGCGGTGGG9CAGAGCCATCTTCTCAAGCCAGATGAAGAGTCTCATGGGCG 3997  
Db 3907 ATGATGAGAGG9CGCGGTGGG9CAGAGCCATCTTCTCAAGCCAGATGAAGAGTCTCATGGGCG 3966  
QY 3998 CGTGCATTCGGGCGGAGAGACAGAGCCGCCACCATCAAAAGACAGACCCAGATTCGGCG 4057  
Db 3967 CGTGCATTCGGGCGGAGAGACAGAGCCGCCACCATCAAAAGACAGACCCAGATTCGGCG 4026  
QY 4058 GGGTTCATCAACAAGGAGATCCCTCGGTCTTAAGTGAAG9CACAAGAGAGCTTAACCTGCGT 4117  
Db 4027 GGGTTCATCAACAAGGAGATCCCTCGGTCTTAAGTGAAG9CACAAGAGAGCTTAACCTGCGT 4086  
QY 4118 CGGAGG9CCAAAGCTCTTAAGG9GAGG9CAGCTCGGCCCCACCGCCTCAAG9AGC 4177  
Db 4087 CGGAGG9CCAAAGCTCTTAAGG9GAGG9CAGCTCGGCCCCACCGCCTCAAG9AGC 4146  
QY 4178 CTGACCGAGG9CTTAAGACGAGG9CGTGGGCGCCCTGAAGCTGAAGCGGCGCCATGAG 4237  
Db 4147 CTGACCGAGG9CTTAAGACGAGG9CGTGGGCGCCCTGAAGCTGAAGCGGCGCCATGAG 4206  
QY 4238 GGGCTGTGTG9CACAAGGTGAAGAGG9CGG9CGCTCATCAATGAAGATCCGCGCAGAGAG 4297  
Db 4207 GGGCTGTGTG9CACAAGGTGAAGAGG9CGG9CGCTCATCAATGAAGATCCGCGCAGAGAG 4266  
QY 4298 CTGCGG9CACAAGCGCCGAGAGCTGCGCCTTGCGCCCGCGGCGCTCAAGAGAGGCTCCATCAAG 4357  
Db 4267 CTGCGG9CACAAGCGCCGAGAGCTGCGCCTTGCGCCCGCGGCGCTCAAGAGAGGCTCCATCAAG 4326  
QY 4358 CAGG9CACAAGCGGCTGAAGTGAAGACAGCGCGG9CTCAACAATG9GTTCCAAAAAGACAGAC 4417  
Db 4327 CAGG9CACAAGCGGCTGAAGTGAAGACAGCGCGG9CTCAACAATG9GTTCCAAAAAGACAGAC 4386  
QY 4418 GTACGCTCCCTCATTCG9AGCCCGG9CGGAGGTTTCCACCCGTCACCCCGCTGATGTG 4477  
Db 4387 GTACGCTCCCTCATTCG9AGCCCGG9CGGAGGTTTCCACCCGTCACCCCGCTGATGTG 4446  
QY 4478 ATGGCCGACGCGCGG9CAGCTG9ACGTGCTGTGTAAGAGAGAGCCTTAAGAGCGCG9CA 4537

Db 4447 ATGGCCGACGCGCGG9CAGCTG9ACGTGCTGTGTAAGAGAGAGCCTTAAGAGCGCG9CA 4506  
QY 4538 GGGACCG9CAGACAGCTCG9GGGG9CTCCATTTGCGGCGG9CGCCGGTCAATTTGTGCTGAG 4597  
Db 4507 GGGACCG9CAGACAGCTCG9GGGG9CTCCATTTGCGGCGG9CGCCGGTCAATTTGTGCTGAG 4566  
QY 4598 CTGG9TAAAGCGCG9CAGAGG9CCCCCTTGAAGCTTAAGAGACACAGG9GACACCTTTTGGCGGC 4657  
Db 4567 CTGG9TAAAGCGCG9CAGAGG9CCCCCTTGAAGCTTAAGAGAGACACAGG9GACACCTTTTGGCGGC 4626  
QY 4658 CACCTCCACAGAGGTGGCGCGGTGAACATG9GGAGG9CAGCGCGCGCTGACAGAGAGGCG 4717  
Db 4627 CACCTCCACAGAGGTGGCGCGGTGAACATG9GGAGG9CAGCGCGCGCTGACAGAGAGGCG 4686  
QY 4718 AGCCTTTCGTCCAGACAGGACCTCCAGAGCCGAAAGCTGAAGTGAAGCGCTGTGAGATC 4777  
Db 4687 AGCCTTTCGTCCAGACAGGACCTCCAGAGCCGAAAGCTGAAGTGAAGCGCTGTGAGATC 4746  
QY 4778 GCCAAGTCCCG9CACAAGCAGCTGTCGGAGACACACCAACCCCATTTGCGCCTTAAGAG 4837  
Db 4747 GCCAAGTCCCG9CACAAGCAGCTGTCGGAGACACACCAACCCCATTTGCGCCTTAAGAG 4806  
QY 4838 CACCTGCTTGGGG9CAGAGG9GGGTG9ACCTGTATG9AGG9CACATCCCGCTGGCCTTC 4897  
Db 4807 CACCTGCTTGGGG9CAGAGG9GGGTG9ACCTGTATG9AGG9CACATCCCTGGCCTTC 4866  
QY 4898 GACCCACCTTCATATACCCCGG9GATCCCTGTGACGAGCGGCTGCTACTACCTGCGCC 4957  
Db 4867 GACCCACCTTCATATACCCCGG9GATCCCTGTGACGAGCGGCTGCTACTACCTGCGCC 4926  
QY 4958 CGACAGCTGG9CCCCAACCCCACTTAAGCCG9CAGCTGTATCCACCTTAATCTATCGGCGCG 5017  
Db 4927 CGACAGCTGG9CCCCAACCCCACTTAAGCCG9CAGCTGTATCCACCTTAATCTATCGGCGCG 4986  
QY 5018 TAACCCGACAGGGG9GCGCTG9AGAACCG9GAGACATCAATGATATCATCATGACCTGCG 5077  
Db 4987 TAACCCGACAGGGG9GCGCTG9AGAACCG9GAGACATCAATGATATCATCATGACCTGCG 5046  
QY 5078 CAGAGATG9CACAACAACAGG9CAGCGCATG9GCCAGCGAGG9GTATGCTGAGG9GCG 5137  
Db 5047 CAGAGATG9CACAACAACAGG9CAGCGCATG9GCCAGCGAGG9GTATGCTGAGG9GCG 5106  
QY 5138 CTTCGCGCGG9AGTCTGCTG9CACTCAACTAAGCTGCGG9TCCCGAGGCTATC 5197  
Db 5107 CTTCGCGCGG9AGTCTGCTG9CACTCAACTAAGCTGCGG9TCCCGAGGCTATC 5166  
QY 5198 GACCTGTCCAAAGT9CACAAGCTGCTG9TGTG9GCGCGG9CAGACAGG9CAGCGGCG 5257  
Db 5167 GACCTGTCCAAAGT9CACAAGCTGCTG9TGTG9GCGCGG9CAGACAGG9CAGCGGCG 5226  
QY 5258 ACCGCAATG9ACCGGCTTGTCTTAAGCTCCCAACCGG9CGCCGCTTCAAGAGCGCGCAC 5317  
Db 5227 ACCGCAATG9ACCGGCTTGTCTTAAGCTCCCAACCGG9CGCCGCTTCAAGAGCGCGCAC 5286  
QY 5318 AGCAGTTCGCCACTTTCGCCAGAGTCCAAACAATTGAACAAACCAACAACAGCTGCC 5377  
Db 5287 AGCAGTTCGCCACTTTCGCCAGAGTTCAAACAATTGAACAAACCAACAACAGCTGCC 5346  
QY 5378 TCGTTCGAGCGG9AGCCAGAACCGG9GATG9AGAGG9G9ACCGG9GATTCGG9AAGCGG9AAG 5437  
Db 5347 TCGTTCGAGCGG9AGCCAGAACCGG9GATG9AGAGG9G9ACCGG9GATTCGG9AAGCGG9AAG 5406  
QY 5438 TCCATTCCTCAAGT9CACAACAAGG9TGAAGCAGCCATCTG9AGAGCTGTGTAAGAG 5497  
Db 5407 TCCATTCCTCAAGT9CACAACAAGG9TGAAGCAGCCATCTG9AGAGCTGTGTAAGAG 5466  
QY 5498 CAGAGCAGCG9CAGAGCG9CAGG9CGG9G9TGG9G9CAGACAGG9CGCGCGCGCC 5557  
Db 5467 CAGAGCAGCG9CAGAGCG9CAGG9CGG9G9TGG9G9CAGACAGG9CGCGCGCGCC 5526  
QY 5558 TCCCACTCCATG9CCACACAGACTGCGCCATCTTCCTTCG9AGCCAGAGATG9CCTTCAG 5617

Db 5527 TCCACTCCCATCCACGACGACTGCCCATCTCCCTTGGAACCGAGATGCCCTTCAG 5586  
Oy 5618 CAGAGACCCAGTGTGCTTCAACAACAGGAGATGATCATCACCGCTGTGAGACC 5677  
Db 5587 CAGAGACCCAGTGTGCTTCAACAACAGGAGATGATCATCACCGCTGTGAGACC 5646  
Oy 5678 AGCAGACCCAGTGTGCTTCAACAACAGGAGATGATCATCACCGCTGTGAGACC 5737  
Db 5647 AGCAGACCCAGTGTGCTTCAACAACAGGAGATGATCATCACCGCTGTGAGACC 5706  
Oy 5738 TTCCTCCCTGACCACTGACCACTGACCACTGACCACTGACCACTGACCACTGACCA 5797  
Db 5707 TTCCTCCCTGACCACTGACCACTGACCACTGACCACTGACCACTGACCACTGACCA 5766  
Oy 5798 ATGAGACCCGCTTGTGCTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5857  
Db 5767 ATGAGACCCGCTTGTGCTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5826  
Oy 5858 GCAAGACCCGACCATGCTTCTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5917  
Db 5827 GCAAGACCCGACCATGCTTCTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5886  
Oy 5918 TCTCTCCCTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5977  
Db 5887 TCTCTCCCTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5946  
Oy 5978 ACCATGACCCGACCATGCTTCTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6037  
Db 5947 ACCATGACCCGACCATGCTTCTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6006  
Oy 6038 GCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6097  
Db 6007 GCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6066  
Oy 6098 ATTCAGAGAACTGGAATCTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6157  
Db 6067 ATTCAGAGAACTGGAATCTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6126  
Oy 6158 GTGAGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6217  
Db 6127 GTGAGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6186  
Oy 6218 CACTGGAAGAGCTGCAACAAGAGCACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6277  
Db 6187 CACTGGAAGAGCTGCAACAAGAGCACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6246  
Oy 6278 CCGCTGAAGCTTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6337  
Db 6247 CCGCTGAAGCTTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6306  
Oy 6338 CAGCCCTGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6397  
Db 6307 CAGCCCTGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6366  
Oy 6398 GTCAACCTGAGAGCAATCAAGTGAAGTATCAACAAGAGTATCAACAAGAGTATCAACA 6457  
Db 6367 GTCAACCTGAGAGCAATCAAGTGAAGTATCAACAAGAGTATCAACAAGAGTATCAACA 6426  
Oy 6458 CAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6517  
Db 6427 CAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6486  
Oy 6518 GTCTGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6577  
Db 6487 GTCTGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6546  
Oy 6578 CCGGACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6637  
Db 6547 CCGGACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6606  
Oy 6638 TCGGCTTGTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6697  
Db 6607 TCGGCTTGTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6666

Oy 6698 GAGCCAGGAGCACTCCCGAGTGTGTATCCCGCTGCTGATCCGGAGTGGGAAACAGACG 6757  
Db 6667 GAGCCAGGAGCACTCCCGAGTGTGTATCCCGCTGCTGATCCGGAGTGGGAAACAGACG 6726  
Oy 6758 GAGCCAGAGAGATGGGCTTCAAGTCTCCAGAGCAACAAGCCAGCCGCTTCTTC 6817  
Db 6727 GAGCCAGAGAGATGGGCTTCAAGTCTCCAGAGCAACAAGCCAGCCGCTTCTTC 6786  
Oy 6818 AGCAAGCTGACCGAGAGCAATCCCGCATGATGATCAAGAGAGAGAGATCAAG 6877  
Db 6787 AGCAAGCTGACCGAGAGCAATCCCGCATGATGATCAAGAGAGAGATCAAG 6846  
Oy 6878 AAGCTGAACCAACAACCGGAATGAGCTGAATTAACAATACAGCAGCTGGGACGAG 6937  
Db 6847 AAGCTGAACCAACAACCGGAATGAGCTGAATTAACAATACAGCAGCTGGGACGAG 6906  
Oy 6938 ATCTTAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6997  
Db 6907 ATCTTAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6966  
Oy 6998 CAGGAACATGCGAGCAACAATGGGGCTGAGAGCAATTAAGAAAGCATATGGGT 7057  
Db 6967 CAGGAACATGCGAGCAACAATGGGGCTGAGAGCAATTAAGAAAGCATATGGGT 7026  
Oy 7058 AATATGACCAAGTGGGAAAGTCCCGGCTGAGGAGCAATCTTTTAACTCTGAAT 7117  
Db 7027 AATATGACCAAGTGGGAAAGTCCCGGCTGAGGAGCAATCTTTTAACTCTGAAT 7086  
Oy 7118 GCCAGTGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7177  
Db 7087 GCCAGTGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7146  
Oy 7178 ACACTCACTCTGCAAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7237  
Db 7147 ACACTCACTCTGCAAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7206  
Oy 7238 AAGGCAAGTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7297  
Db 7207 AAGGCAAGTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7266  
Oy 7298 GTGCACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7357  
Db 7267 GTGCACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7326  
Oy 7358 AAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7417  
Db 7327 AAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7386  
Oy 7418 GCGGATGTCATGCTTCCCAACCCCAACCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 7477  
Db 7387 GCGGATGTCATGCTTCCCAACCCCAACCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 7446  
Oy 7478 GGGCCCAACCAAGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7537  
Db 7447 GGGCCCAACCAAGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7506  
Oy 7538 CTCTCCGACAGGAGTGA 7555  
Db 7507 CTCTCCGACAGGAGTGA 7524

RESULT 15  
ACA62250 standard, cDNA, 7521 BP.  
XX ACA62250;  
XX AC  
XX AC  
XX DT  
XX 12-AUG-2003 (first entry)  
XX Human nuclear receptor corepressor SMRPe coding region cDNA.  
XX Human nuclear receptor corepressor SMRPe; gene therapy;  
XX Human; ss; gene; SMRPe; nuclear receptor corepressor; gene therapy;  
KW



[illegible]

D	b	2407	GTGTCCTCCCAAGAGAGAGAAAGAGAGAGAGACCGACACAGCGCCCACTGTGAGAGAGAGG	2466
O	y	2522	GAGAGACAGAAAGCCCCCGCGCTTGAGAGCTGGCACTGACACAGAGAAAGCCGAGAG	2581
D	b	2467	GAGAGACAGAAAGCCCCCGCGCTGAGAGCTGGCACTGTGACACAGAGAAAGGCGAGAG	2526
O	y	2582	CCCTTCAAAGACGAGTACACGAGAGAAAGCCGAGAGAGGGGCGCGCAAGGGGCAAGAGAGCG	2641
D	b	2527	CCCTTCAAAGACGAGTACACGAGAGAAAGCCGAGAGAGGGGCGCGCAAGGGGCAAGAGAGCG	2586
O	y	2642	GAGGCGCTGAGGCGCACGCGCCAGAGGGGCGCTCAAGGCAAGAAAGAGAGAGGGGAGG	2701
D	b	2587	GAGGCGCTGAGGCGCACGCGCCAGAGGGGCGCTCAAGGCAAGAAAGAGAGGGGAGG	2646
O	y	2702	GGCAGGGGCAACCACTGCGCAAGAGCTCTGGGCGCCCCCAGAGACAGCGACTCTCAGTCTACC	2761
D	b	2647	GGCAGGGGCAACCAAGGCAAGAGCTCTGGGCGCCCCCAGAGACAGCGACTCTCAGTCTACC	2706
O	y	2762	TGCAGTGCAGACGAGGTGATGAGGCGGAGGCGGCAAGAAACCGCTGTCTCTCCCA	2766
D	b	2707	TGCAGTGCAGACGAGGTGATGAGGCGGAGGCGGCAAGAAACCGCTGTCTCTCCCA	2766
O	y	2822	AGGCGCGAGCCCTCTCAACCCCGACTGGAGACCCCGGGGCAATGGCTCAACCCAGAAAGCA	2881
D	b	2767	AGGCGCGAGCCCTCTCTCAACCCCGACTGGAGACCCCGGGGCAATGGCTCAACCCAGAAAGCA	2826
O	y	2882	CTGAGACTGAAAGCAGCTGAAAGCAGCGAGCGGCTGCATCCCGCCCATCGAGTCAACCAA	2941
D	b	2827	CTGAGACTGAAAGCAGCTGAAAGCAGCGAGCGGCTGCATCCCGCCCATCGAGTCAACCAA	2886
O	y	2942	GTCCTATAGGCCCCCGGGAGAGCGAGCTTCCACCAAGCCAGCTCCCGAGCCCCACCG	3001
D	b	2887	GTCCTATAGGCCCCCGGGAGAGCGAGCTTCCACCAAGCCAGCTCCCGAGCCCCACCG	2946
O	y	3002	CCACCGCAAAACCTGACGCGCGGAGAGAGGCGCCCTAGAGAGCTGGACAGAGCGCCCGG	3061
D	b	2947	CCACCGCAAAACCTGACGCGCGGAGAGAGGCGCCCTAGAGAGCTGGACAGAGCGCCCGG	3006
O	y	3062	GGCAAGACAGAGAGCCCGGAGACCCCGCGGACAGAGAG-----3100	3100
D	b	3007	GGCAAGACAGAGAGCCCGGAGACCCCGCGGACAGAGAG-----3100	3066
O	y	3101	---GCCTTGGCAGCGGAGGCGCCAGAAAGCTGCTGGAGACCCCTTGTGACTTCCGGC	3157
D	b	3067	CCAGCTTTCGACGAGCGGAGGCGCCAGAAAGCTGCTGGAGAGCCCCCTTGTGACTTCCGGC	3126
O	y	3158	CTGCGCTTCCCCCGGTGCGCCCCCGGTGAGGTATCAAGGCTTCCCGGACTTGGCGGAGACCC	3217
D	b	3127	CTGCGCTTCCCCCGGTGCGCCCCCGGTGAGGTATCAAGGCTTCCCGGACTTGGCGGAGACCC	3186
O	y	3218	TCAAGCTTTCCTTACGCTCCACCGTGGTCAACCTGCGCTCGAGGAGCTTCCAGACTGAGC	3277
D	b	3187	TCAAGCTTTCCTTACGCTCCACCGTGGTCAACCTGAGTCAACCTGCGCTCGAGACTGAGC	3246
O	y	3278	CGGCGCGTCTGTGCGCGGCGCCAGCCACCATCTCCAAACCGGCTCCCTCATCTCTTGC	3337
D	b	3247	CGGCGCGTCTGTGCGCGGCGCCAGCCACCATCTCCAAACCGGCTCCCTCATCTCTTGC	3306
O	y	3338	AAGCACCCCAAGCGCTCTCGAGAGGCAAAATGAGTGCATCTTCCAGGAATGTGGTCTAG	3397
D	b	3307	AAGCACCCCAAGCGCTCTCGAGAGGCAAAATGAGTGCATCTTCCAGGAATGTGGTCTAG	3366
O	y	3398	CTTCAAGTCCCGTATCTCAGAGCAATGCAAGAGCCCGGTGGGCGCTGTCACTATGGGGCTG	3457
D	b	3367	CTTCAAGTCCCGTATCTCAGAGCAATGCAAGAGCCCGGTGGGCGCTGTCACTATGGGGCTG	3426
O	y	3458	CCCCTGCGCATGAGACCCCAAAAAGCTTGCACCTTCAAGCGAGTGAAGCAGAGCAGCTG	3517
D	b	3427	CCCCTGCGCATGAGACCCCAAAAAGCTTGCACCTTCAAGCGAGTGAAGCAGAGCAGCTG	3486
O	y	3518	TCCCCACAGGGGCGAGGCTTGGGCGCACCGGAGAGGCTTGGGGGTGCCACAGCCCGAGAGGCG	3577
D	b	3487	TCCCCACAGGGGCGAGGCTTGGGCGCACCGGAGAGGCTTGGGGGTGCCACAGCCCGAGAGGCG	3546



QY 3578 TCCGCTGAGAGGGACAGCTCTGGGCTCAAGTTCCGGCGGAAAGCATCACCAAAGCAATT 3637  
Db 3547 TCCGCTGAGAGGGACAGCTCTGGGCTCAAGTTCCGGCGGAAAGCATCACCAAAGCAATT 3606  
QY 3638 CCCAGCAACAGGGTGTCCCTCGGACAGCGCATCACTACCCGGGCTCCATCCACCGAC 3697  
Db 3607 CCCAGCAACAGGGTGTCCCTCGGACAGCGCATCACTACCCGGGCTCCATCCACCGAC 3666  
QY 3698 ACGCAGCTGAAGCTCTGTATCAAGGGACCATCAACAGATATCATGGGACAGAGCCG 3757  
Db 3667 ACGCAGCTGAAGCTCTGTATCAAGGGACCATCAACAGATATCATGGGACAGAGCCG 3726  
QY 3758 AGTCGCTTGAACCGCGGCGCGGAGGACAGCCCTGCCAAGGGCCACGTCACTTACGAAGC 3817  
Db 3727 AGTCGCTTGAACCGCGGCGCGGAGGACAGCCCTGCCAAGGGCCACGTCACTTACGAAGC 3786  
QY 3818 AAGAAAGGGCCAGCTTTGTCTTATGAGGGTGAGATGTCTGTACCCAGTGCTCCAAAGG 3877  
Db 3787 AAGAAAGGGCCAGCTTTGTCTTATGAGGGTGAGATGTCTGTACCCAGTGCTCCAAAGG 3846  
QY 3878 GACGCGAGAAAGAGCTCAAGAACCCGCCCATGAGACGGCGCGGCCCAAGGGCCACTATGAC 3937  
Db 3847 GACGCGAGAAAGAGCTCAAGAACCCGCCCATGAGACGGCGCGGCCCAAGGGCCACTATGAC 3906  
QY 3938 ATGATGAGAGGGCCGCGTGGGACAGACCACTTCTCAGCCAGCATCGAAAGTCTATGGAC 3997  
Db 3907 ATGATGAGAGGGCCGCGTGGGACAGACCACTTCTCAGCCAGCATCGAAAGTCTATGGAC 3966  
QY 3998 CGTGCATTCGCCCGCGGACGACACGCCCCCACTTCAAAAGACAGACCACTATCCG 4057  
Db 3967 CGTGCATTCGCCCGCGGACGACACGCCCCCACTTCAAAAGACAGACCACTATCCG 4026  
QY 4058 GGGTCCATCAACAAGGATCCCTCGGTCTTACGTGAGGACAGAGAGACTACCTGCGT 4117  
Db 4027 GGGTCCATCAACAAGGATCCCTCGGTCTTACGTGAGGACAGAGAGACTACCTGCGT 4086  
QY 4118 CGGAGAGCCAAAGCTCTAAAGCGGAGGACAGCTTCGCGCCCAACGCGCTCACGGGAC 4177  
Db 4087 CGGAGAGCCAAAGCTCTAAAGCGGAGGACAGCTTCGCGCCCAACGCGCTCACGGGAC 4146  
QY 4178 CTGACCGAGGCTTAACAACGAGGCTCTGGGCCCCCTGAAAGTGAAGCCGCGCATGAG 4237  
Db 4147 CTGACCGAGGCTTAACAACGAGGCTCTGGGCCCCCTGAAAGTGAAGCCGCGCATGAG 4206  
QY 4238 GACCCTGATGGCCACGCTGAAGAGGCGGAGCGCTCATCATGAGATCCCGCGGAGAG 4297  
Db 4207 GACCCTGATGGCCACGCTGAAGAGGCGGAGCGCTCATCATGAGATCCCGCGGAGAG 4266  
QY 4298 CTGCGGCAACGCGCGGAGCTGCGCTGCGCGCGGCGCTCAAGAGAGCTCATCAAG 4357  
Db 4267 CTGCGGCAACGCGCGGAGCTGCGCTGCGCGCGGCGCTCAAGAGAGCTCATCAAG 4326  
QY 4358 CAGGGACACCCGCTCAAGTACGACACCGGCGGCTCACTGAGTCCAAAAAGCAAGAC 4417  
Db 4327 CAGGGACACCCGCTCAAGTACGACACCGGCGGCTCACTGAGTCCAAAAAGCAAGAC 4386  
QY 4418 GTACGCTCCCTCATATGAGGAGCCCGGCGGAGCTTCCACCCGCTGATGATG 4477  
Db 4387 GTACGCTCCCTCATATGAGGAGCCCGGCGGAGCTTCCACCCGCTGATGATG 4446  
QY 4478 ATGCGCGACGCGCGGAGCACTGGAAGCTGCGCTGCTACGAGAGAGCTGTAAGAGCGGCA 4537  
Db 4447 ATGCGCGACGCGCGGAGCACTGGAAGCTGCGCTGCTACGAGAGAGCTGTAAGAGCGGCA 4506  
QY 4538 GGGACCGCGACGAGCTCGGGGAGGCTCAATTGCGCGCGGCGCGCTCATTTGTGCTGAG 4597  
Db 4507 GGGACCGCGACGAGCTCGGGGAGGCTCAATTGCGCGCGGCGCGCTCATTTGTGCTGAG 4566  
QY 4598 CTGGGTAAAGCGCGGACAGACCCCTGAACCTATGAGAGCAACAGGGGCAACCTTTGCGGCG 4657  
Db 4567 CTGGGTAAAGCGCGGACAGACCCCTGAACCTATGAGAGCAACAGGGGCAACCTTTGCGGCG 4626

QY 4658 CACCTCCACGAGGTTGCGCCGCTGACATCGCGGAGCCCAACGCGCGCTTCAGAGAGGC 4717  
Db 4627 CACCTCCACGAGGTTGCGCCGCTGACATCGCGGAGCCCAACGCGCGCTTCAGAGAGGC 4686  
QY 4718 AGCCTTTGTCCAGAAAGGCAATCCAGAGACCGAAAGCTGAAGCTGAGGCTGTGAGATC 4777  
Db 4687 AGCCTTTGTCCAGAAAGGCAATCCAGAGACCGAAAGCTGAAGCTGAGGCTGTGAGATC 4746  
QY 4778 GCCAAGTCCCGCACAGACCGGTGCGGAGACACCCACACCCCATCTGCGCTTATGAG 4837  
Db 4747 GCCAAGTCCCGCACAGACCGGTGCGGAGACACCCACACCCCATCTGCGCTTATGAG 4806  
QY 4838 CACTGTCTTGGGGGTGAGTGGCGTGAACCTGTATGCGAGCCACATCCCTGCGCTTC 4897  
Db 4807 CACTGTCTTGGGGGTGAGTGGCGTGAACCTGTATGCGAGCCACATCCCTGCGCTTC 4866  
QY 4898 GACCCCACTTCATACCCCGGAGCATCCCTGTGAGCGAGCGGCTGACTACCTGCGC 4957  
Db 4867 GACCCCACTTCATACCCCGGAGCATCCCTGTGAGCGAGCGGCTGACTACCTGCGC 4926  
QY 4958 CGACACTGGGCCCCCAACCCCACTTACCCGACCTGTACCCACCTTACCTTACCGCGGC 5017  
Db 4927 CGACACTGGGCCCCCAACCCCACTTACCCGACCTGTACCCACCTTACCTTACCGCGGC 4986  
QY 5018 TACCCCGACAGCGCGGCTGTGAGAACCGGACACATCATCATGACTTACATCACTCG 5077  
Db 4987 TACCCCGACAGCGCGGCTGTGAGAACCGGACACATCATCATGACTTACATCACTCG 5046  
QY 5078 CAGCAGATGCAACCAACAGGCGCAACGCGCATGCGCGCAGCGAGTATGCTGAGGGGC 5137  
Db 5047 CAGCAGATGCAACCAACAGGCGCAACGCGCATGCGCGCAGCGAGTATGCTGAGGGGC 5106  
QY 5138 CTCTGCGCCCGAGTCTCGCTGAGCACTCAACTACGCTGCGGAGTCCCGAGGCACTATC 5197  
Db 5107 CTCTGCGCCCGAGTCTCGCTGAGCACTCAACTACGCTGCGGAGTCCCGAGGCACTATC 5166  
QY 5198 GACTGTTCCTCAAGTGCACACTGTCTGTGCTCGTCCCGGACACAGGCACTCCAGCC 5257  
Db 5167 GACTGTTCCTCAAGTGCACACTGTCTGTGCTCGTCCCGGACACAGGCACTCCAGCC 5226  
QY 5258 ACCGCAATGAGCCGCTTGCTGCTACCTCCCAACCGGCGCCAGCCCTTGAAGGCGCGCAC 5317  
Db 5227 ACCGCAATGAGCCGCTTGCTGCTACCTCCCAACCGGCGCCAGCCCTTGAAGGCGCGCAC 5286  
QY 5318 AGCAGCTCCCACTCTCCCGAGAGTCCAAACAATTGACAAAACAAACACACAGCTCC 5377  
Db 5287 AGCAGCTCCCACTCTCCCGAGAGTCCAAACAATTGACAAAACAAACACACAGCTCC 5346  
QY 5378 TCGTCCGAGCGGAGCGAGACCGGGATTCGAGAGCGGAGCCGGGATCGGGAGGAAAAAG 5437  
Db 5347 TCGTCCGAGCGGAGCGAGACCGGGATTCGAGAGCGGAGCCGGGATCGGGAGGAAAAAG 5406  
QY 5438 TCCATTCCTCAGTCCACAGACAGCGGTGAGACGACCATCTGAGACTGTGTACAGAG 5497  
Db 5407 TCCATTCCTCAGTCCACAGACAGCGGTGAGACGACCATCTGAGACTGTGTACAGAG 5466  
QY 5498 CAGAGCAGCGCAGCAGCGGACGAGCGCGGAGGTGAGGAGCAAGCAGCGCGCCCGCC 5557  
Db 5467 CAGAGCAGCGCAGCAGCGGAGCGGACGCGCGGAGGTGAGGAGCAAGCAGCGCGCCCGCC 5526  
QY 5558 TCCCATCTCCATGCCCACAGACACTCGCCCATCTTCCCTGAGACCCAGATGCTTCAG 5617  
Db 5527 TCCCATCTCCATGCCCACAGACACTCGCCCATCTTCCCTGAGACCCAGATGCTTCAG 5586  
QY 5618 CAGAGACCAATGAGCTTCAACAACAGGCAATGAAGGTAATCATACCGCTGTGAGAGCC 5677  
Db 5587 CAGAGACCAATGAGCTTCAACAACAGGCAATGAAGGTAATCATACCGCTGTGAGAGCC 5646  
QY 5678 AGCAAAGCCACGAGTCTGAGGTCCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 5737  
Db 5647 AGCAAAGCCACGAGTCTGAGGTCCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 5706  
QY 5738 TTCCCACTTGCCACCACTGCGCACCTGAGCGGAGCACTTCGATGAGGCTTACCTTACCTTC 5797

```

Db      ||||| 5707 TTCCCACTGCAACCCACTGACCCTGGAGGCACTCGATGGGTCTACCTACCTTC 5766
Qy      ||||| 5798 ATGGAGCCCGCTCTTGTCGCCCAAGAGAGGCCCCCGGGTGGCCCGGCAAGCGGCCCCCA 5857
Db      ||||| 5767 ATGGAGCCCGCTCTTGTCGCCCAAGAGAGGCCCCCGGGTGGCCCGGCAAGCGGCCCCCA 5826
Qy      ||||| 5858 GAGACAACCGGCAATGCTTCTCTGCGCAAGCCCCGAGCCCGCTCGGGGTGAGAGCCGCG 5917
Db      ||||| 5827 GAGAGACCGGCAATGCTTCTCTGCGCAAGCCCCGAGCCCGCTCGGGGTGAGAGCCGCG 5886
Qy      ||||| 5918 TCTCTCCCGAGAAAGGGCTCGAGAGCCCGGCTTGTGTCTCTGCGCAAGCC 5977
Db      ||||| 5887 TCTCTCCCGAGAAAGGGCTCGAGAGCCCGGCTTGTGTCTCTGCGCAAGCC 5946
Qy      ||||| 5978 ACCATGCCCCGCAACCCCTGCGAAAGAACTTCGACCTTCACACGCGAGCCCGGAGCCCGCG 6037
Db      ||||| 5947 ACCATGCCCCGCAACCCCTGCGAAAGAACTTCGACCTTCACACGCGAGCCCGGAGCCCGCG 6006
Qy      ||||| 6038 GCGCCACTGCGCTCGGCTCGAGACCGGCAACCGGAAAGAACTCAAGTAAACCTTTTCC 6097
Db      ||||| 6007 GCGCCACTGCGCTCGGCTCGAGACCGGCAACCGGAAAGAACTCAAGTAAACCTTTTCC 6066
Qy      ||||| 6098 ATCCAGGAATGGAATCTCCGTTCTTGAGTTACACGCGAGAGCTTACAGCCCGAGAGG 6157
Db      ||||| 6067 ATCCAGGAATGGAATCTCCGTTCTTGAGTTACACGCGAGAGCTTACAGCCCGAGAGG 6126
Qy      ||||| 6158 GTGAGAGCCCGTCAAGCCCTGTGAGCTCAACCAAGTTCAGACCAAGAGGCTCCCCAG 6217
Db      ||||| 6127 GTGAGAGCCCGTCAAGCCCTGTGAGCTCAACCAAGTTCAGACCAAGAGGCTCCCCAG 6186
Qy      ||||| 6218 CACTGGAAGAGTGCAGAAAGGCCACTGAGAGGGGAGCTGCGGCCCAAGGAGCCAGG 6277
Db      ||||| 6187 CACTGGAAGAGTGCAGAAAGGCCACTGAGAGGGGAGCTGCGGCCCAAGGAGCCAGG 6246
Qy      ||||| 6278 CCGGTGAAGCTTGGCGGGAGGCGGCCACTTCCCAACCTGCGCGCTGCTGAGAGC 6337
Db      ||||| 6247 CCGGTGAAGCTTGGCGGGAGGCGGCCACTTCCCAACCTGCGCGCTGCTGAGAGC 6306
Qy      ||||| 6338 CAGCCCTCGTCCAGCGCCGCTGCTCCAGACCGGCCCGAGGGGTCAAGAGTCAACAGCGGGT 6397
Db      ||||| 6307 CAGCCCTCGTCCAGCGCCGCTGCTCCAGACCGGCCCGAGGGGTCAAGAGTCAACAGCGGGT 6366
Qy      ||||| 6398 GTCAACCTGGGCCAGGACATCATGTAGGTTCATCAACAGGACTACACCGGCAACCA 6457
Db      ||||| 6367 GTCAACCTGGGCCAGGACATCATGTAGGTTCATCAACAGGACTACACCGGCAACCA 6426
Qy      ||||| 6458 CAGCAGCTCAGGCAACCCCTGCGCGGCCCTCTTACTCTTCCCTGCGGCGCAGCTGCCCC 6517
Db      ||||| 6427 CAGCAGCTCAGGCAACCCCTGCGCGGCCCTCTTACTCTTCCCTGCGGCGCAGCTGCCCC 6486
Qy      ||||| 6518 GTCTGTGACCTTCGCGCGGCCCAACAGTGACTTCACTCCCGGCCGAGCAATGTGTC 6577
Db      ||||| 6487 GTCTGTGACCTTCGCGCGGCCCAACAGTGACTTCACTCCCGGCCGAGCAATGTGTC 6546
Qy      ||||| 6578 CCGGCGCGGTGCTCCCGCAACGCGGAAGGGGGCAAGAGTTCAGAGCCAAACAAGAG 6637
Db      ||||| 6547 CCGGCGCGGTGCTCCCGCAACGCGGAAGGGGGCAAGAGTTCAGAGCCAAACAAGAG 6606
Qy      ||||| 6638 TCGGTCTTGGGTGTGTGTGAGGACGGTATTTGAACCTGTGTCCCAACGAGGGGCAATGAC 6697
Db      ||||| 6607 TCGGTCTTGGGTGTGTGTGAGGACGGTATTTGAACCTGTGTCCCAACGAGGGGCAATGAC 6666
Qy      ||||| 6698 GAGCCAGGCACTCCCGGAGTGTGTATCCGCTGTGTACCGGGATGGGGAACAGACG 6757
Db      ||||| 6667 GAGCCAGGCACTCCCGGAGTGTGTATCCGCTGTGTACCGGGATGGGGAACAGACG 6726
Qy      ||||| 6758 GAGCCAGGCAATGGGCTCAAGTCTCAGGCAACAGGCAAGCGGCAAGCTTCTTC 6817
Db      ||||| 6727 GAGCCAGGCAATGGGCTCAAGTCTCAGGCAACAGGCAAGCGGCAAGCTTCTTC 6786
Qy      ||||| 6818 AGCAAGCTGACGAGGCAACTCCGCAATGTCAAGTCCAGAGCAAGAGATCAACAAG 6877

```

```

Db      ||||| 6877 AGCAAGCTGACCGAGAGCAATCCGCAATGTCAAGTCCAGAGCAAGAGATCAACAAG 6846
Qy      ||||| 6878 AAGCTGAACCCCAACACCGGAATGAGCCTGAATACATATACAGGCTGGAGAGAG 6937
Db      ||||| 6847 AAGCTGAACCCCAACACCGGAATGAGCCTGAATACATATACAGGCTGGAGAGAG 6906
Qy      ||||| 6938 ATCTTCAATATGCCCCGCAATCAACCGGAACAGGCTTATGACCTATGAAAGCCAGCGGT 6997
Db      ||||| 6907 ATCTTCAATATGCCCCGCAATCAACCGGAACAGGCTTATGACCTATGAAAGCCAGCGGT 6966
Qy      ||||| 6998 CAGGAACATGCAACCAACATGGGAGTGAAGGCCATATTAAGAGCACTATGGT 7057
Db      ||||| 6967 CAGGAACATGCAACCAACATGGGAGTGAAGGCCATATTAAGAGCACTATGGT 7026
Qy      ||||| 7058 AATATGACAGTGGGAAGAGTCCCGCGCTCAGGCGCAAGGCTTTTAACTCTGAAT 7117
Db      ||||| 7027 AATATGACAGTGGGAAGAGTCCCGCGCTCAGGCGCAAGGCTTTTAACTCTGAAT 7086
Qy      ||||| 7118 GCGAGTGCAGGCTGCCCGCTGTATGACCATTAACCGCTGACGAGCGAGTGAACA 7177
Db      ||||| 7087 GCGAGTGCAGGCTGCCCGCTGTATGACCATTAACCGCTGACGAGCGAGTGAACA 7146
Qy      ||||| 7178 AACTCACTCGCAGGTGGCGGCGGAAGGCCAAGTCTTGGCAGACCCAGCGCGA 7237
Db      ||||| 7147 AACTCACTCGCAGGTGGCGGCGGAAGGCCAAGTCTTGGCAGACCCAGCGCGA 7206
Qy      ||||| 7238 AAGCCAAATCCCGGCGCGCGGCGCTGAGATCTGGGGAACCGGCCAACCCTGTCTCTCA 7297
Db      ||||| 7207 AAGCCAAATCCCGGCGCGCGGCGCTGAGATCTGGGGAACCGGCCAACCCTGTCTCTCA 7266
Qy      ||||| 7298 GTGCACTGGAAGGAGACTGCAACCGCGGAGCGGCTCAACCAACCGGCTGTGGAGAG 7357
Db      ||||| 7267 GTGCACTGGAAGGAGACTGCAACCGCGGAGCGGCTCAACCAACCGGCTGTGGAGAG 7326
Qy      ||||| 7358 AGGCTCTGTCGCAAGTTCACGCGCAATCCCTTCAACCCCTGTATCATGGGCTGCA 7417
Db      ||||| 7327 AGGCTCTGTCGCAAGTTCACGCGCAATCCCTTCAACCCCTGTATCATGGGCTGCA 7386
Qy      ||||| 7418 GGGGTGTCAATGGCTTTCCCAACCCCAACGGGCTTCCCGCGGCGAGCGGGCCCTCGCT 7477
Db      ||||| 7387 GGGGTGTCAATGGCTTTCCCAACCCCAACGGGCTTCCCGCGGCGAGCGGGCCCTCGCT 7446
Qy      ||||| 7478 GGGCCCAACAGGCTGGGACGAGAGCCCAAGCACTGCTGTGCACTACAGACA 7537
Db      ||||| 7447 GGGCCCAACAGGCTGGGACGAGAGCCCAAGCACTGCTGTGCACTACAGACA 7506
Qy      ||||| 7538 CTCTCGACAGCGAG 7552
Db      ||||| 7507 CTCTCGACAGCGAG 7521

```

Search completed: March 12, 2006, 06:25:27  
 Job time : 3076 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2006, 12:21:21 ; Search time 2098 Seconds  
(without alignments)  
9426.045 Million cell updates/sec

Title: US-09-522-753-4  
Perfect score: 8564.8  
Sequence: 1 catgctgggcccacacagc.....caaaaaaaaaaaaaaaaa 8571

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7673375 seqs, 115364844 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA New:  
1: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	630	7.4	7940	9 US-11-245-147-191	Sequence 191, App
2	190.6	2.2	559	12 US-11-136-527-1693	Sequence 1693, App
3	190.6	2.2	559	12 US-11-136-527-1693	Sequence 1693, App
4	186.2	2.2	609	6 US-09-925-065A-16738	Sequence 841377, App
5	156.8	1.8	421	6 US-09-925-065A-16738	Sequence 516005, App
6	126.2	1.5	600	8 US-10-750-623-1127	Sequence 1127, App
7	126.2	1.5	600	8 US-10-750-623-1127	Sequence 1127, App
8	115	1.3	30191	7 US-10-330-773-631	Sequence 631, App
9	113.4	1.3	72352	7 US-10-330-773-631	Sequence 631, App
10	112.4	1.3	302	6 US-09-925-065A-16738	Sequence 516003, App
11	98.6	1.2	153376	12 US-11-121-086-5	Sequence 5, App11
12	84	1.0	1061	6 US-09-925-065A-16738	Sequence 720226, App
13	83.6	1.0	409	6 US-09-925-065A-16738	Sequence 16738, App
14	82.4	1.0	2093	9 US-11-072-512-1666	Sequence 1666, App
15	81.2	0.9	1400	12 US-11-136-527-1693	Sequence 3376, App
16	81.2	0.9	2280	12 US-11-136-527-1693	Sequence 3376, App
17	79	0.9	204603	7 US-10-330-773-631	Sequence 325, App
18	76	0.9	1728	8 US-10-750-623-1127	Sequence 42639, App
19	76	0.9	1728	8 US-10-750-623-1127	Sequence 42639, App
20	75.8	0.9	118864	7 US-10-330-773-826	Sequence 826, App

21	74.8	0.9	2539	12 US-11-136-527-215	Sequence 215, App
22	73.4	0.9	26214	7 US-10-330-773-222	Sequence 222, App
23	73	0.9	1027	12 US-11-136-527-94	Sequence 94, App1
24	71.8	0.8	2183	9 US-11-072-512-1064	Sequence 1064, App
25	71.6	0.8	2479	12 US-11-136-527-3303	Sequence 3303, App
26	71.6	0.8	9474	12 US-11-052-554A-526	Sequence 526, App
27	71.6	0.8	194186	7 US-10-330-773-395	Sequence 395, App
28	70.8	0.8	7474	12 US-11-069-834-49	Sequence 49, App1
29	70.6	0.8	600	12 US-11-136-527-7399	Sequence 7399, App
30	70.2	0.8	610	6 US-09-925-065A-16738	Sequence 780159, App
31	70.2	0.8	610	6 US-09-925-065A-16738	Sequence 780160, App
32	69.2	0.8	365	12 US-11-043-752-3306	Sequence 3306, App
33	69.2	0.8	365	12 US-11-043-752-3306	Sequence 3306, App
34	68.6	0.8	4146	12 US-11-052-554A-522	Sequence 522, App
35	68.2	0.8	657	6 US-09-925-065A-16738	Sequence 952675, App
36	67.8	0.8	1619	8 US-10-821-234-10	Sequence 10, App1
37	67.8	0.8	2772	12 US-11-052-554A-531	Sequence 531, App
38	67.6	0.8	26667	8 US-10-995-561-13375	Sequence 13375, App
39	67.4	0.8	244	12 US-11-043-752-3912	Sequence 3912, App
40	66.4	0.8	207908	12 US-11-112-908-19	Sequence 21, App1
41	66.4	0.8	212805	12 US-11-112-908-19	Sequence 19, App1
42	66.2	0.8	5706	12 US-11-052-554A-519	Sequence 519, App
43	66	0.8	649	6 US-09-925-065A-16738	Sequence 228, App
44	66	0.8	38920	7 US-10-330-773-228	Sequence 228, App
45	65.8	0.8	5706	12 US-11-052-554A-519	Sequence 519, App

ALIGNMENTS

RESULT 1  
US-11-245-147-191  
Sequence 191, Application US/11245147  
Publicat ion No. US20060030541A1  
GENERAL INFORMATION:  
APPLICANT: GARCIA, TERESA  
APPLICANT: ROMAN ROMAN, SERGIO  
APPLICANT: BARON, ROLAND  
APPLICANT: CALL, KATHERINE  
APPLICANT: THEILHABER, JOACHIM  
APPLICANT: CONNOLLY, TIMOTHY  
APPLICANT: JACKSON, AMANDA  
APPLICANT: BUSHNELL, STEVEN  
APPLICANT: RAMADI, GEORGES  
TITLE OF INVENTION: GENES INVOLVED IN OSTEOGENESIS, AND METHODS OF USE  
FILE REFERENCE: 37991-0023  
CURRENT FILING DATE: 2005-10-07  
PRIORITY FILING DATE: 2005-10-07  
PRIORITY FILING DATE: 2002-04-05  
PRIORITY FILING DATE: 2002-04-05  
PRIORITY FILING DATE: 2001-04-05  
NUMBER OF SEQ ID NOS: 246  
SOFTWARE: Patent In Ver. 3.2  
SEQ ID NO 191  
LENGTH: 7940  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Homo sapiens nuclear receptor co-repressor 1  
OTHER INFORMATION: (NCOR1), mRNA  
US-11-245-147-191

Query Match 7.4%; Score 630; DB 9; Length 7940;  
Best Local Similarity 65.1%; Pred. No. 4.8e-115;  
Matches 1015; Conservative 0; Mismatches 495; Indels 48; Gaps 4;

CGCGACGCGTCTCCAGAGAGCTGATCCAGACATGACCGGTGGACGAGAGATC 550  
CTTCAAAACCTCTCAAGAGAGCTGATTAATACAGACTATGAGATCGTGTGATGAGAAATT 816  
ACCATGCTAGAGACAGACATCTCTAAGCTGAAGAGAGCAAGCAGCTGAGAGAGG 610

Db	817	GC AAAAGT AAGAA CAG CAGATCTCTTAACTG AAAAAAAGAA CAA CAA CAGCTTGAAGAAAG	876
Qy	611	GCTGCCAAGCCGCCGAGCTTGAGAAAGCCGTGTACCGCGCCCATGAGTCGAAAGAC	670
Db	877	GCACCTAAACCTCCTGAGCTGAGAAAGCCGCTGTCCCTCCTCTGTGTGAGCAGAAACAC	936
Qy	671	CGCAGCCCTGGTGCAGATCATCTACGACGAAACCGGAAAGAGCTGAAGCTGCACATCGG	730
Db	937	CGCAGTATTTGTC A AATTAATTAATGATGAGATCGGAAAAAAGCAGAAAGATCTCATAAA	996
Qy	731	ATTTCGAAAGGCTCTGGGGCCCAAGGTGAGCTGCGCTGTACAACCAACCTCCCGCAGCC	790
Db	997	ATTTTGAAGGCTCTTGCCCAAAAGTTGAATCTGCCACTGTATTAACAGCCCATCAGATACC	1056
Qy	791	CGCAGTATTCATGAGAACATCAAAATAAACAGCGATGCGGAAAGACTATCTTGATAC	850
Db	1057	AAGGTGTACATGAGAACATCAAGCAAAACAGGTGATGAGGAAAAAACTCATTTATTTT	1116
Qy	851	TTTCAAGAGGAGGATTCACGCTTCGAAAACATATGAAAGCGAAAGTTCTGCGACGCTATGAC	910
Db	1117	TTTAAAAAATAATCATGCAAGAAAAAATAAGGAAACAAAAATCTGCGACGTTATGAT	1176
Qy	911	CAGCTCATGAGGCTCTTGAAAAAAGGTGGAGCGCATCGAAAAACACCGCGCCGCGCG	970
Db	1177	CAGCTCATGAGGCGATGAGGAAAAAAGTGGACAAATATGAAATATCTCTCGAGGAAA	1236
Qy	971	GCCAAAGAGCAAGTGTCGCGAGTACTACAAAAAGCAGTTCCCTGAGATCGCAAGCAG	1030
Db	1237	GCTAAAGAAACAAACAAAGGAAATACTATGAAAAGCAGTTTCCAGAAATTCGAAAAACA	1296
Qy	1031	CGCGAGCTGCAGAGAGCGCATGCAAGCAGGCTGGCCACGCGGGCAGTGGCTGTCCATG	1090
Db	1297	AGAAACACGCAAGAAAGATTTTCAG---CGATTTGGCGAGAGGGAGCTGTCTTTCAGCC	1353
Qy	1091	TTCGCGCGCCGCGAGCAGAGCAGAGGTGCAGAGATCATCGATGGCTCTCAGAGAGGAG	1150
Db	1354	ACCATTTGCTAGAGATGACATGAGATTTCTGAATTTATGATGGCTCTCTGAGCAGGAG	1413
Qy	1151	AACCTGAGGAAGCAGATGCGCAGCTGGCCGTGATCCCGCCATGCTGTACGAGCCTGAC	1210
Db	1414	AATATATGAGAAACAAATTCGGCAGCTCTGTGTGATTCACCTATGATGTTTGAATGCAAA	1473
Qy	1211	CAGCAGCGCATCAAGTTTCATCAACATGAAACGGCTTATGSCCGAACCCCATGAAGGTGAC	1270
Db	1474	CAAAAGCAGAGTCAAGTTTCATTAACATGATGGCTTATGAGAGCACCTATGAAGTGTAT	1533
Qy	1271	AAAAACCGCCAGGTCATACGACATGTGAGTGTGACGAGAGAGAGAGACCTTCCGGAGAG	1330
Db	1534	AAATATATGAGCAGTTTATGATGTTTGAACCTACCAATGAAGAAAGAGATCTTTAAGGACAG	1593
Qy	1331	TTTCATGCGCATCCCAAGAACCTTGGGCTGTATGCGCATATTCCTGAGAGGAAAGCAGTG	1390
Db	1594	TTTATTCAGCATCCAAAAAACTTTGAGCTAATTTGCTATATCTTGTGAGAGAAAGTGT	1653
Qy	1391	GCTGATGCGTCTCTATTACTACCTGACTAAGAAATGAGAACTATTAAGCGCTGTG	1450
Db	1654	CCTGATGTTGTTTGTATTACTAATTAACAAAGAAAAATGAGAAATTAAGCTCCTGCTC	1713
Qy	1451	AGAGCGAGCTATCGGCGCGCGCGCGCAGAGGCGCAGAGCAACAAAGCAGCAGCAGCAGCAG	1510
Db	1714	AGAAAGAAATTAATGAGAAACGCAAGGCGCAAAACAGCA-----AATTTCTGACCTT	1765
Qy	1511	CAGACGACGACGACGACGACCCATGCCCCGACGACGCGAGGAGGAAAGATGAGAG	1570
Db	1766	CGCAAGAAAGAAATGATGAGAAAGAAAGAAAGATTAAGCGAAAAAACAAGAAAAAAG	1825
Qy	1571	GAGAAAGAAAGAGAGCGGAGAGAGAGAGAGAACCGGAGGTGTGAGAAACGACAAAGGA	1630
Db	1826	AAGAAAGAAA-----GAATGATGAGAGGAAAAAAGATGAAAAAAGAAAGCTCCAAAGAA	1878
Qy	1631	GACCTCTCAAGAGAGAGACAGACGACCTTCAGGGGAGGACAAAGCATGAGAGAGGCTT	1690
Db	1879	AATTCACAGAGAAAGGACAGATATGATGTATCAGCAGAAAGAACTGAGGAAAGAGAGCA	1938

QY	1691	GTGCGCTTCGAAAGCGCCGAAATCTGCCAACGCCAGGAGAGCCGAAAGGCCGATCAC	1750
Db	1939	GCCACACCCCGGGGGCGGAAGACTGCCAACAGTCAGGGCCGCTAAGGGCGCGATCAC	1998
QY	1751	CGCTCAATGGCTAATGA-----GGCAACAGCGAG	1780
Db	1999	AGGTTCATGAAACGAAAGCTGCGAGCTGCCAGTGTCTGACGCCGACCGCTACTGAAAG	2058
QY	1781	GAGGCGCATCAACCCCCAGCAGAGCGCGAGCTGGCTCTCATGAGCTGAATGAGAGTTCT	1840
Db	2059	CCCCACCACTCTGTGCCACCCGACAGAAACCATTTCTACAGAGCGCTGTGAGACCTCT	2118
QY	1841	CGTTGACAGAAAGAAATGAGAAACAGCCAGAAAGGTCTCTTGAAACACGCGCCGAC	1900
Db	2119	CGATGGACAGAAAGAAATGAGAGTTGCTAAAGAGGTCTAGTAGAACAATGTCGTAC	2178
QY	1901	TGTGCGCGCAATCGCCCGAGTGTGTGGCTCCAGACTGTGTGCGAGTGAACAATTCTAC	1966
Db	2179	TGGGACAGCAATGTGTAATATGTGTGGAAACGAAAGTAGAAGTCAATGTAATTAATCTTAT	2238
QY	1961	TTCACTACTAAGAAAGAGGACAGAACTCGATGAGATCTTGACAGACACAACTGAAG	2018
Db	2239	TTTAACATATAAGGCGACACAACTTGAACAACCTTTACAGCAGCATTAACAGAAAA	2296
RESULT 2			
US-11-136-527-1693/c			
; Sequence 1693, Application US/11136527			
; Publication No. US20050287570A1			
; GENERAL INFORMATION:			
; APPLICANT: Wyeth			
; APPLICANT: Mounts, William M			
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes			
; FILE REFERENCE: 031896-041000 (AM101086)			
; CURRENT APPLICATION NUMBER: US/11/136,527			
; CURRENT FILING DATE: 2005-05-25			
; PRIOR APPLICATION NUMBER: US 60/574,294			
; PRIOR FILING DATE: 2005-05-26			
; NUMBER OF SEQ ID NOS: 362830			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 1693			
; LENGTH: 559			
; TYPE: DNA			
; ORGANISM: Rattus norvegicus			
; FEATURES:			
; NAME/KEY: misc_feature			
; LOCATION: (543)..(543)			
; OTHER INFORMATION: n is a, c, g, or t			
US-11-136-527-1693			
Query Match			
Best Local Similarity 68.6%; Pred. No. 2,4e-28; DB 12; Length 559;			
Matches 395; Conservative 20; Mismatches 135; Indels 26; Gaps 11;			
QY	7997	CCGCTTACCGCTCTGACAGATGAGCCAGGC-CCTGTACAGCCCCAGTGCCTCTTC	8055
Db	555	CCCGCTGGCGCTNTGACAGACAGACGTTCCAGCTTATCCTGCCCATGCTGTATCG	496
QY	8056	CGGTCCCAAGACTGCCCCAGCCACGACGATGCTGGAACCAAGTCAGGCCAGGTGG	8115
Db	495	CAGTCCCAAGGCGACCCAGCCACAR-----GACTGGAGCCCATCRGACGAGTGG	441
QY	8116	CGACCAAAAGGCGCAGGTGGCGCTGGGGGGAACGATGCTCCGAGGACTGATGTTTT	8175
Db	440	TGCAC-AAAGGCGCGGATAGGCT--GGGMAAGATGTTTGACGAGAACCGACGTTG	384
QY	8176	TTTACACATCTGTTGCGCGACGGGTGGAGAAAGACAGATGTAATGATGTGTGTT	8235
Db	383	ATTCCCGGCTGTGATATGCGAT--GGCGGGAAGGACGTTAATG--GTGTGCT	327
QY	8236	TACAGGATATATTTTGTATACCTTCAATGAATTAATTCAATGTTTTACGACGAAAGAA	8295



Db 326 WACAGGATATTTTGTGTAACCTTCAATGATTAATGATGCTTCAACAAGAGGA 267  
QY 8296 CTACCCAGATATTAAGTCTGTGCTTTGATCTGTCTTACCGTTCAAGAGCGGTG 8355  
Db 266 CTCACCC-GTGTACTGCGCTGTGC-TTGGTMTTGTCTTGTCCGAGGACATG 209  
QY 8356 CAGGCGGACAGTGTGATCCCATCACTCGCAGAGCCAGGAGGCGGAGCTGTCTGCA 8415  
Db 208 CCGGCCA--AGGCGGTGCGCCGCACTACGAGACTGAGGAGGCAAGGAGGCGGCA--- 154  
QY 8416 CCGCCCGCTGTGTCTTCCCTCCCTCCCTCTCTGCGGAGAAATGATGCTATTC 8475  
Db 153 --CAGGCACTGTGCGCTCCCTCCCTCTCTGCGGAGAAATGATGCTATTC 96  
QY 8476 GTGCGCGCATTTGGCAGAGGTGTGATTTCTGTCAATTACACAGTGTCTTAATTA 8535  
Db 95 GTGCGCGCTCTGTGC-----ATGCGGCAATTTCTGTCAATTACAGTGTCTCAATTA 41  
QY 8536 AAGCGAATTATTAATCTCCAAAAA 8571  
Db 40 AARGCAATTATTAATCTCCARMMMAA 5  
RESULT 3  
US-11-136-527-5789  
; Sequence 5789, Application US/11136527  
; Publication No. US20050287570A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William M  
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes  
; FILE REFERENCE: 031896-04100 (AM101086)  
; CURRENT APPLICATION NUMBER: US/11/136,527  
; PRIOR FILING DATE: 2005-05-25  
; PRIOR APPLICATION NUMBER: US 60/574,294  
; PRIOR FILING DATE: 2005-05-26  
; NUMBER OF SEQ ID NOS: 362830  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5789  
; LENGTH: 559  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (17)-(17)  
; OTHER INFORMATION: n is a, c, g, or t  
US-11-136-527-5789  
Query Match 2.2%; Score 190.6; DB 12; Length 559;  
Best Local Similarity 68.6%; Pred. No. 2.4e-28;  
Matches 395; Conservative 20; Mismatches 135; Indels 26; Gaps 11;  
QY 7997 CCGCTTAGCGCTCTGTGACAGATGAGCGAGGC-CCTGTCCAGCCCAAGTGCCTGTT 8055  
Db 5 CCGCGTCGCGCTTGTGACAGACAGAGCTTCCAGCTTATCTGCCCCATGCTGTATCG 64  
QY 8056 CCGTCCCAACAGCTGCCCCAGCAAGAGATGCTGGAACCAATCAAGCCAGCGTGG 8115  
Db 65 CAGTCCCAAGAGCCCAAGCCCAAGCAAR-----GACTGGAGCCATCTGAGCCAGGTGG 119  
QY 8116 CCGAAGAAAGGCGAGTGTGCGCTGAGGAGAAAGATGCTCCGAGACTGCACTTTT 8175  
Db 120 TGCAC-AAAAGGCGCGGTATGCGCT--GGGAAAGGATGTTTGCAAGAACCGGACGTT 176  
QY 8176 TTTCAACATCTGTGCGGAGCGGTGGAGAAAGGAGAGATGTAATGATGTGTGTT 8235  
Db 177 ATTCCCGGCTGTGCTATGAGCAT-GGCGGAGAAAGCAGGTATGATG--GTGTGCT 233  
QY 8236 TACAGGATATTTTGTGATCTTCAATTAATTAATTAATTAATTAATTAATTAATTA 8295  
Db 234 WACAGGATATTTTGTGATCTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 293  
QY 8296 CTACCCAGATATTAAGTCTGTGCTTTGATCTGTCTTACCGTTCAAGAGCGGTG 8355

Db 294 CTCACCC-GTGTACTGCGCTGTGC-TTGGTMTTGTCTTGTCCAGAGCAATG 351  
QY 8356 CAGGCGGACAGTGTGATCCCATCACTCGCAGAGCCAGGAGGCGGAGCTGTCTGCA 8415  
Db 352 CCGGCCA--AGGCGGTGCGCCGCACTACGAGACTGAGGAGGCAAGGAGGCGGCA--- 406  
QY 8416 CCGCCCGCTGTGTCTTCCCTCCCTCCCTCTCTGCGGAGAAATGATGCTATTC 8475  
Db 407 --CAGGCACTGTGCGCTCCCTCCCTCTCTGCGGAGAAATGATGCTATTC 464  
QY 8476 GTGCGCGCATTTGGCAGAGGTGTGATTTCTGTCAATTACACAGTGTCTTAATTA 8535  
Db 465 GTGCGCGCTCTGTGC-----ATGCGGCAATTTCTGTCAATTACAGTGTCTCAATTA 519  
QY 8536 AAGCGAATTATTAATCTCCAAAAA 8571  
Db 520 AARGCAATTATTAATCTCCARMMMAA 555  
RESULT 4  
US-09-925-065A-841377/C  
; Sequence 841377, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 841377  
; LENGTH: 609  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-841377  
Query Match 2.2%; Score 186.2; DB 6; Length 609;  
Best Local Similarity 98.4%; Pred. No. 1.8e-27;  
Matches 188; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 7185 CCTGCGCAGTGTGCGGAGAGCCCAAGTCTGTGCGAGACCCAGACCGAAAAAGCA 7244  
Db 212 CCGCTGCAAGTGTGCGGAGAGCCCAAGTCTGTGCGAGACCCAGACCGAAAAAGCA 153  
QY 7245 AGTCCCGGCGCGGCTGTGCGATCTGAGGAGCCGACACCTGTGTCTCTCAATGCACT 7304  
Db 152 AGTCCCGGCGCGGCTGTGCGATCTGAGGAGCCGACACCTGTGTCTCTCAATGCACT 93  
QY 7305 CCGAGGAGACTGCAACCGCGGAGCGGCTCAACCAACCGGAGTGTGAGGAGACAGGCGCT 7364  
Db 92 CCGAGGAGACTGCAACCGCGGAGCGGCTCAACCAACCGGAGTGTGAGGAGACAGGCGCT 33  
QY 7365 CGTCCGAGGT 7375  
Db 32 CGTCCGAGGT 22  
RESULT 5  
US-09-925-065A-516005  
; Sequence 516005, Application US/09925065A

```

Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIORITY APPLICATION NUMBER: US 60/243,096
PRIORITY FILING DATE: 2000-10-24
PRIORITY APPLICATION NUMBER: US 60/252,147
PRIORITY FILING DATE: 2000-11-20
PRIORITY APPLICATION NUMBER: US 60/250,092
PRIORITY FILING DATE: 2000-11-30
PRIORITY APPLICATION NUMBER: US 60/261,766
PRIORITY FILING DATE: 2001-01-16
PRIORITY APPLICATION NUMBER: US 60/289,846
PRIORITY FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 516005
LENGTH: 421
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-516005

```

	Query Match	Best Local Similarity	1.8% Matches	Score 156.8	DB 6	Length 422
			95.8%	Pred. No. 1.1e-21		
			161	Conservative	0	Mismatches 7
						Indels 0
						Gaps 0
QY	5182	TCGCCGAGGCATCATCGACTCTGTCCCAAGTGCACACCTGTGCTGTCTGTGTCGCCCCAC	5241			
Db	228	TCCCGCAGGCATCATCGACTCTGTCCCAAGTGCACACCTGTGCTGTGTCGCCCCAC	287			
QY	5542	ACCGGCAACCCCAAGCCACCGCCATGGAACCGCTTGTCTTAACCTCCCAACCGGCCCCACCC	5301			
Db	288	ACCGGCAACCCCAAGCCACCGCCATGGAACCGCTTGTCTTAACCTCCCAACCGGCCCCACCC	347			
QY	5302	CTTGAGCAGCGCCGACAGCAGCTCCCACTCTCCCAAGAGTCCAC	5349			
Db	348	CTTGAGCAGCGCCGACAGCAGCTCCCACTCTCCCAAGAGTCCAC	395			

```

RESULT 6
US-10-750-185-1127/c
; Sequence 1127, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1127
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMET13672
; US-10-750-185-1127

```

Query Match	1.5%;	Score 126.2;	DB 8;	Length 600;
Best Local	Similarly	91.2%;	Pred. No. 1.3e-15;	
Matches 134;	Conservative	0;	Mismatches 13;	Indels 0;
				Gaps 0;

[illegible]

```

RESULT 7
US-10-750-623-1127/c
; Sequence 1127, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-1
; CURRENT APPLICATION NUMBER: US/10/750.623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1127
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMET13672
US-10-750-623-1127

```

	Query Match	1.5% Best Local Similarity	Score 91.2%	DB Pred. No.	Length 1.3e-15	Matches 134	Conservative 0	Indels 13	Gaps 0
Oy	1498 GCACGACAGCAGCAGCAGCAGCAGCAGCCCATGCCCGCAGCAGCAGCAGGA	1557							
Dd	196 GCACGACAGCAGCAGCAGCAGCAGCAGCAGCAGATGCCCGTGACGCCAGAGA	137							
Oy	1558 GAAGATGAGAGGAAAGCAAAGCGGAGAGAGAGAAAGCCGAGGTGG	1617							
Dd	136 GAAAGATGAGAGGAGAAAGGAGGTGGAGAGAGAGAGAACCCGACGTGA	77							
Oy	1618 GAAGCAAGGAAAGCCCTCCCAAGGA	1644							
Dd	76 GAATGACAAGAGGAACTGATCAAGTA	50							

```

RESULT 8
US-10-330-773-631/C
; Sequence 631, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 631
; LENGTH: 30191
; TYPE: DNA
; ORGANISM: Mus musculus

```



Qy 2055 AAGCCCGCGCGGCGCAGGAGGCTGATTCCTCCCGCTGTGTGAGATGAGAGA 2114  
Db 63034 AGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 62975  
Qy 2115 TCGAGGCGCTCGGCGCTGAGCGGAAATGAGAGAGATGTGTGAGAGGCTGA 2166  
Db 62974 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 62923

## RESULT 10

US-09-925-065A-516003/C  
; Sequence 516003, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 516003  
; LENGTH: 302  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-516003

Query Match 1.3%; Score 112.4; DB 6; Length 302;  
Best Local Similarity 99.1%; Pred. No. 6.3e-13;  
Matches 113; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2214 GCCCAGCACTGTCAACAACAGCTCAGACCCGAGAGCATCCCTCTCTCTCACTGAGG 2273  
Db 114 GCACAGCCACTGTCAACAACAGCTCAGACCCGAGAGCATCCCTCTCTCTCACTGAGG 55  
Qy 2274 CCGCCAGGACACAGGCGAGATGGGCCCAAGCCCCCAGCCACCTTGGGCGCCG 2327  
Db 54 CCGCCAGGACACAGGCGAGATGGGCCCAAGCCCCCAGCCACCTTGGGCGCCG 1

## RESULT 11

US-11-121-086-5  
; Sequence 5, Application US/1121086  
; Publication No. US20050266459A1  
; GENERAL INFORMATION:  
; APPLICANT: NIELSEN, TIM S.  
; APPLICANT: POULSEN, TIM S.  
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES  
; FILE REFERENCE: 09138.6000-00000  
; CURRENT APPLICATION NUMBER: US/11/121,086  
; CURRENT FILING DATE: 2005-05-04  
; PRIOR APPLICATION NUMBER: 60/567,570  
; PRIOR FILING DATE: 2004-05-04  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 5  
; LENGTH: 153376  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-121-086-5

Query Match 1.2%; Score 98.6; DB 12; Length 153376;

Best Local Similarity 44.9%; Pred. No. 7.8e-10;  
Matches 577; Conservative 0; Mismatches 699; Indels 9; Gaps 5;

Qy 1417 GACTTAAGAAAGATGAACTTAAGAGCTGTGTGAGACGAGCTTATCGGCGCCGCGGAA 1476  
Db 15043 GAGGAGGAGGAGAAACAGCTGTAGACACTGTGGGACCTCTCCAGATGCCGGGCACTG 15102  
Qy 1477 GAGCCAGGAGCAACAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1536  
Db 15103 GAGACAGAGAGGAGGAGGAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15162  
Qy 1537 GCCCCGACGAGCCAGGAGGAGAAAGATGAGAAAGAGAGAGAGAGAGAGAGAGAG 1596  
Db 15163 GAGAGAACAGAGGAGGAGGAGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15222  
Qy 1597 GAGAGAGAGCCGAGGAGTGGAGAAACGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1656  
Db 15223 GAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15282  
Qy 1657 CACCTCAGGAGGAGCAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1715  
Db 15283 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15342  
Qy 1716 CCAACAGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1775  
Db 15343 GAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15402  
Qy 1776 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1834  
Db 15403 GAGACATCCCAAGTGCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15462  
Qy 1835 AGTTCTCGCTGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1894  
Db 15463 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15522  
Qy 1895 GCGAATGAGTGGCCATG-CCCCGATGTGGGCTCCAGACTGTGTGAGTGAAGA 1953  
Db 15523 CCAGATGCGGCGGCGCTGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15582  
Qy 1954 CTTTCACTTCAACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2013  
Db 15583 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15642  
Qy 2014 GAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2073  
Db 15643 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15702  
Qy 2074 CGAGAGGCTGCAATTCCTGCGCGGTGTGAGAGATGAGAGATGAGAGCTGTGGGCTGAG 2133  
Db 15703 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15762  
Qy 2134 CGGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2193  
Db 15763 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15822  
Qy 2194 GCCCAGAGGAGAGATGAGTGGCCCGCAGCACTGTCAACAACAGCTCAGACCCGAGAGAT 2253  
Db 15823 GGTGCTTAGACACTGTGGGAGCTGTCCAGATGTCCAGGGGCACTGTGAGACAGAGAGAGAG- 15881  
Qy 2254 CCGCTCTCTCACTAGGCGCGCCAGAGACACAGGAGAGAGAGAGAGAGAGAGAGAG 2313  
Db 15882 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15941  
Qy 2314 CACCTGGGCGCCAGCGGGCCACCCCGAGGCGCACCCACCCACAGAGAGAGAGAGAG 2373  
Db 15942 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 16001  
Qy 2374 GAGCCGCAATTAGAGCCACCCGCGCTTGAAGCCACCCGAGAGAGAGAGAGAGAGAG 2433  
Db 16002 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 16061  
Qy 2434 ACCCCATGCGCTCTGCACTCTCTGTGTGTCCCAAGAGAGAGAGAGAGAGAGAGAG 2493





/ CURRENT APPLICATION NUMBER: US/11/072,512  
/ CURRENT FILING DATE: 2005-03-07  
/ PRIOR APPLICATION NUMBER: US 60/350,978  
/ PRIOR FILING DATE: 2002-01-25  
/ PRIOR APPLICATION NUMBER: JP 2001-379298  
/ PRIOR FILING DATE: 2001-11-05  
/ NUMBER OF SEQ ID NOS: 4096  
/ SOFTWARE: PatentIn Ver. 2.1  
/ SEQ ID NO 1666  
/ LENGTH: 2093  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-11-072-512-1666

Query Match 1.0%; Score 82.4; DB 9; Length 2093;  
Best Local Similarity 47.1%; Pred. No. 6.9e-07;  
Matches 356; Conservative 0; Mismatches 391; Indels 9; Gaps 3;  
QY 936 AGTGTGAGCGCATTCGAAACCAACCCGCGCGGCGCCAGAGAGCAAGGTGCGCGAGT 995  
Db 997 AGGAGGAGAGATGTGTGAGACAGAGAAAGCGCTCGAGAGCAGAGAGAGAGCGCGG 1056  
QY 996 ACTCGAAAGAGTTCCTGATCCGAGACAGCGAGCTGAGAGCGCATGCG 1055  
Db 1057 AGCAGAGAGAGAGCTGCGGAGCAGAAAGCTCGGAGCAGAGAGCAGATGCGAG 1116  
QY 1056 GCAGGTTGGCGCGCGGCGAGTGGGCTGTCCATG--TCGCGCGCGCCGAGCGACAG 1113  
Db 1117 AGCAGAGAGAGAGATGTGTGAGCAGAGAGAGAAATGCGGAGCAGAGAGAGATGT 1176  
QY 1114 GGTGTCAAGATCATCATGTCCTCTCA-GAGCAGAGAACTTGAGAAAGCATGCGCC 1172  
Db 1177 GAGAGCAGAGAGAGCGCTGTGTGAGCAGAGAAACAGATGCGGAGCAGAGCAGAA 1236  
QY 1173 AGCTGGCGCGTATCCCGCCCATCTGTGACGACGTCGACCAAGAGCGCATCAATTCA 1232  
Db 1237 TCGCGGACCAAGAGAGAGATGTGTGAGCAGAGCAGAGGCTTACGGAGAAAGAG 1296  
QY 1233 ACATGAACGGGCTTATGGCCGACCCCATGAGGTGTAACAAGCCGAGGTCAATGA 1292  
Db 1297 GGATCGGAGAGAGAGAAATGTGTGAGCAGGTGGAAGATGTGCGGAGAGAAAG 1356  
QY 1293 TGTGAGTGAAGAGAGAGAGACCTTTCGGGAGAAATTCAATGACATCCCAAG 1352  
Db 1357 TGCAGAGACAGAGAGAAAGACGCGGACAGAGAGAGAAATGCAAGAGAGAGAG 1416  
QY 1353 TTGGCTGTGATCATTCCTGTGAGAGAGAGACATGTGCTGATGCGCTCTATTACT 1412  
Db 1417 TACGGAGCGGAGAA-----GAGATGCGGAAAGAGAGAGACGATGCGGAGCAG 1470  
QY 1413 ACTTGACTTAAGAAATGAGAACTTAAGAGCTGTGAGACGAGCTATCGCGCGCG 1472  
Db 1471 AGGAGAAATGTGAGAGAGAGAGAGAAATATGTGTGAGAGAGAGAAAGAGTGGCAG 1530  
QY 1473 GCAAGAGCCAGCAGCAACAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1532  
Db 1531 AGCAGAGGCTGCGGAAACAAGAGAGAGCTGTGGAACAGGAGAAAGATGCAAGCAG 1590  
QY 1533 CCATGCCCCGAGCAGCAGAGAGAGAAAGATGAGAGAGAGAGAGAGAGAGAGAG 1592  
Db 1591 AGGAGAAATTTGGAGCAGAGAGAGAGAGATACGGGACAGAGAGAGATGTGGGGCAG 1650  
QY 1593 AGAGAGAGAGAGCCGAGGTGTGAGAAAGCAAGAGAAAGCTCTCAAGAGAGAGAG 1652  
Db 1651 AGAGAGAAATGTGTGAGAGAGAGAGAGATGTGTGGGGCAGAGAGAGATGTGAGAGAG 1710  
QY 1653 ACGACACTCAGGGGAGAGCAACGACGAGAGAGAG 1688  
Db 1711 AGAGAGAGATACGAGATCAGAAAGAGAGATCAGG 1746

RESULT 15  
US-11-136-527-7472/c

/ Sequence 7472, Application US/1136527  
/ Publication No. US20050287570A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Wyeth  
/ APPLICANT: Mounts, William M  
/ TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes  
/ FILE REFERENCE: 031896-041000 (AM101086)  
/ CURRENT APPLICATION NUMBER: US/11/136,527  
/ CURRENT FILING DATE: 2005-05-25  
/ PRIOR APPLICATION NUMBER: US 60/574,294  
/ PRIOR FILING DATE: 2005-05-26  
/ NUMBER OF SEQ ID NOS: 362830  
/ SOFTWARE: PatentIn version 3.2  
/ SEQ ID NO 7472  
/ LENGTH: 1400  
/ TYPE: DNA  
/ ORGANISM: Rattus norvegicus  
US-11-136-527-7472

Query Match 0.9%; Score 81.2; DB 12; Length 1400;  
Best Local Similarity 59.6%; Pred. No. 1.1e-06;  
Matches 137; Conservative 0; Mismatches 93; Indels 0; Gaps 0;  
QY 1413 ACTTGACTTAAGAAATGAGAACTTAAAGCTGTGTGAGACGAGCTATCGCGCGCG 1472  
Db 507 AGCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 448  
QY 1473 GCAAGAGCCAGCAGCAACAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1532  
Db 447 AGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 388  
QY 1533 CCATGCCCCGAGCAGCAGAGAGAGAAAGTGAAGAGAGAGAGAGAGAGAGAGAG 1592  
Db 387 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 328  
QY 1593 AGAGAGAGAGAGCCGAGGTGTGAGAAAGCAAGAGAGAGCTCTCAAG 1642  
Db 327 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 278

Search completed: March 12, 2006, 15:36:41  
Job time : 2101 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2006, 10:17:44 ; Search time 5962 Seconds  
(without alignments)  
11888.098 Million cell updates/sec

Title: US-09-522-753-4  
Perfect score: 8564.8  
Sequence: 1 catgctgcgcacacagc.....caaaaaaaaaaaaaa 8571

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 10%  
Listing first 45 summaries

Database : Published Applications NA Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8561	99.7	8561	6 US-10-174-014-11	Sequence 11, App1
2	8541	99.7	8561	10 US-11-052-344-1	Sequence 1, App1
3	8459.2	98.8	8667	5 US-10-087-192-653	Sequence 653, App
4	8324.6	97.2	8686	3 US-09-819-104A-1	Sequence 1, App1
5	8324.6	97.2	8686	6 US-10-174-014-4	Sequence 4, App1
6	8324.6	97.2	8686	8 US-10-723-860-1739	Sequence 1739, App
7	8222.2	96.0	9079	8 US-10-723-860-6114	Sequence 6114, App
8	7554	88.2	7554	6 US-10-414-692-14	Sequence 14, App1
9	7334.6	85.6	7521	3 US-09-819-104A-3	Sequence 3, App1
10	5037.2	58.8	8544	3 US-09-819-104A-4	Sequence 4, App1
11	5037.2	58.8	8544	9 US-10-764-420-1526	Sequence 1526, App
12	4797	56.0	7386	3 US-09-819-104A-6	Sequence 6, App1
13	4543	53.0	7534	5 US-10-087-192-650	Sequence 650, App
14	2618.6	30.6	2930	5 US-10-146-473-14	Sequence 14, App1
15	2618.6	30.6	2930	6 US-10-174-014-13	Sequence 13, App1
16	1138.6	13.3	23380	5 US-10-087-192-652	Sequence 652, App
17	1132.4	13.2	221000	6 US-10-174-014-12	Sequence 12, App1
18	949.4	11.1	1027	3 US-09-925-297-9	Sequence 9, App1
19	898.4	10.5	956	3 US-09-887-527-42	Sequence 42, App1
20	898.4	10.5	956	7 US-10-796-174-42	Sequence 42, App1
21	630	7.4	7940	6 US-10-341-434-235	Sequence 235, App
22	630	7.4	7940	8 US-10-473-974-191	Sequence 191, App
23	630	7.4	7940	9 US-10-887-553A-664	Sequence 664, App

24	630	7.4	7940	9 US-10-956-157-1580	Sequence 1580, App
25	630	7.4	7940	9 US-10-745-237-221	Sequence 221, App
26	626.8	7.3	2745	5 US-10-414-692-15	Sequence 15, App1
27	553.6	6.5	692	5 US-10-027-632-102554	Sequence 102554, App
28	553.6	6.5	692	5 US-10-027-632-102555	Sequence 102555, App
29	553.6	6.5	692	6 US-10-027-632-102554	Sequence 102554, App
30	553.6	6.5	692	6 US-10-027-632-102555	Sequence 102555, App
31	542.4	6.3	718	8 US-10-278-698-515	Sequence 515, App
32	542.4	6.3	718	8 US-10-278-698-1031	Sequence 1031, App
33	432.8	5.1	9141	5 US-10-087-192-649	Sequence 649, App
34	409.4	4.8	427	9 US-10-779-543-12895	Sequence 12895, App
35	376	4.4	872	9 US-10-779-543-8757	Sequence 8757, App
36	370	4.3	381	3 US-09-918-995-7653	Sequence 7653, App
37	314.4	3.7	458	3 US-09-918-995-20064	Sequence 20064, App
38	268.6	3.1	520	3 US-09-920-300A-485	Sequence 485, App
39	268.6	3.1	520	5 US-10-033-528-485	Sequence 485, App
40	268.6	3.1	520	6 US-10-099-926-485	Sequence 485, App
41	268.6	3.1	520	6 US-10-961-527-485	Sequence 485, App
42	258.4	3.0	3969	6 US-10-006-285-471	Sequence 471, App
43	211.4	2.5	499	3 US-09-918-995-19696	Sequence 19696, App
44	191.4	2.2	390	6 US-10-355-716-86	Sequence 86, App1
45	187.8	2.2	789	9 US-10-779-543-5226	Sequence 5226, App

## ALIGNMENTS

RESULT 1  
US-10-174-014-11  
; Sequence 11, Application US/10174014  
; Publication No. US20040005292A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Susan M. Freiler  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION  
; FILE REFERENCE: PTS-0012  
; CURRENT APPLICATION NUMBER: US/10/174,014  
; NUMBER OF SEQ ID NOS: 73  
; SEQ ID NO 11  
; LENGTH: 8561  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)...(7555)  
US-10-174-014-11

Query Match Best Local Similarity 99.7%; Pred. No. 0; Length 8561;  
Matches 8560; Conservative 1; Mismatches 0; Indels 10; Gaps 1;

QY	1	CATGTCGGGCTCCACACAGCTTGTGGACAGAGCTGGAGGGCCACTAGAGCCCGCTAACCC	60
DB	1	CATGTCGGGCTCCACACAGCTTGTGGACAGAGCTGGAGGGCCACTAGAGCCCGCTAACCC	60
QY	61	GGCCCAAGCCCTTCTTACCCAGTGCAGATCGCCGAGACACAGGAGTGGGGCTCCT	120
DB	61	GGCCCAAGCCCTTCTTACCCAGTGCAGATCGCCGAGACACAGGAGTGGGGCTCCT	120
QY	121	GGAGTACGACGACCACTCCCGGAGATGCTCCCACTGTGGCGGGGCTCCATCA	180
DB	121	GGAGTACGACGACCACTCCCGGAGATGCTCCCACTGTGGCGGGGCTCCATCA	180
QY	181	GGCCCAAGCCCGGAGGCTCCCTGCTGTGAGTTCAGCCCGGAGATGAAGGTTCCCA	240
DB	181	GGCCCAAGCCCGGAGGCTCCCTGCTGTGAGTTCAGCCCGGAGATGAAGGTTCCCA	240
QY	241	GGAGTTCACCTGCGGACGAGTCCCATATACCTGCGGAGCTGGGAAATGACAGAT	300
DB	241	GGAGTTCACCTGCGGACGAGTCCCATATACCTGCGGAGCTGGGAAATGACAGAT	300

QY 301 GAGATTCAATGAAGCAAGCGCTTGCGCTAGAGCTGCTGACCCCTGCTGCGACC 360  
Db 301 GGAAGTTCAATGAAACCAAGCGCCTTGCGCTAGAGCTGCTGACCCCTGCTGCGACC 360  
QY 361 GTCAACCCCTGCTGCGCAACCGGCGCAAGCTTGCGGAGATCTGAAGAATTCAACCAAGAACCTTAG 420  
Db 361 GTCAACCCCTGCTGCGCAACCGGCGCAAGCTTGCGGAGATCTGAAGAATTCAACCAAGAACCTTAG 420  
QY 421 CTTGACCGGCAAGCTGGAACCGGCTGTCTCCCGACGCCCCCGCACATCGAACCTTGAGCT 480  
Db 421 CTTGACCGGCAAGCTGGAACCGGCTGTCTCCCGACGCCCCCGCACATCGAACCTTGAGCT 480  
QY 481 GGAAGCTGTGCTCGGCAACCGGCTGTCCAAGAGAGAGCTGATCCAGAACATGGAACCGGCTGGA 540  
Db 481 GGAAGCTGTGCTCGGCAACCGGCTGTCCAAGAGAGAGCTGATCCAGAACATGGAACCGGCTGGA 540  
QY 541 CCGAAGATCAACCAATGATAGACACAGATCTCTAAGCTGAAGAAAGAACCAACAGCT 600  
Db 541 CCGAAGATCAACCAATGATAGACACAGATCTCTAAGCTGAAGAAAGAACCAACAGCT 600  
QY 601 GAGAGAGAGAGCTGCGCAAGCGCGCGAGCTGGAAGCCGCTGTGACCGCGCGCATCGA 660  
Db 601 GAGAGAGAGAGCTGCGCAAGCGCGCGAGCTGGAAGCCGCTGTGACCGCGCGCATCGA 660  
QY 661 GTCGAAGCACCGCAGCTGTGTGCAAGATCATCGAAGAACCGGAGAGAGCTGGAAGC 720  
Db 661 GTCGAAGCACCGCAGCTGTGTGCAAGATCATCGAAGAACCGGAGAGAGCTGGAAGC 720  
QY 721 TGCAATGCGGATTTGGAAGGCTGTGGGGCCCAAGGTGAGCTGCGCTGTACAACAGGC 780  
Db 721 TGCAATGCGGATTTGGAAGGCTGTGGGGCCCAAGGTGAGCTGCGCTGTACAACAGGC 780  
QY 781 CTCCGACACCCCGCAGATATCATGAGAACATCAAAATTAACCAAGCGATGCGAAGAGCT 840  
Db 781 CTCCGACACCCCGCAGATATCATGAGAACATCAAAATTAACCAAGCGATGCGAAGAGCT 840  
QY 841 AATCTTTGATCTTCAAGAGAGAGATCAAGCTGTGAAAACATGGAAGCAGAGATTCTGCCA 900  
Db 841 AATCTTTGATCTTCAAGAGAGAGATCAAGCTGTGAAAACATGGAAGCAGAGATTCTGCCA 900  
QY 901 GCGCTATGACCAAGCTCATGGAAGCGCTTGGAATAAAAGGTGAGCGCATGGAACCAACC 960  
Db 901 GCGCTATGACCAAGCTCATGGAAGCGCTTGGAATAAAAGGTGAGCGCATGGAACCAACC 960  
QY 961 GCGCGCGCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
Db 961 GCGCGCGCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
QY 1021 CCGCAAGCACCGCAGCTGTGAGAGCGCATGCAAGACAGGCTGGGCGACGCGGCGAGTGG 1080  
Db 1021 CCGCAAGCACCGCAGCTGTGAGAGCGCATGCAAGACAGGCTGGGCGACGCGGCGAGTGG 1080  
QY 1081 GCTGTCCATGTGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140  
Db 1081 GCTGTCCATGTGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140  
QY 1141 AGAGCAGAGAGAACTGAGAGAACAGATGCGCAGCTGCGCTGATCCCGCATCTGTA 1200  
Db 1141 AGAGCAGAGAGAACTGAGAGAACAGATGCGCAGCTGCGCTGATCCCGCATCTGTA 1200  
QY 1201 CGACGCTGACCAAGACGCGCATCAAGTTCAATGAACGCGGCTTTATGGCCGACCCCAT 1260  
Db 1201 CGACGCTGACCAAGACGCGCATCAAGTTCAATGAACGCGGCTTTATGGCCGACCCCAT 1260  
QY 1261 GAAAGTTGTAACAAGACCGCAGAGTCAATGAACATGTGAGAGAGAGAGAGAGAGAGAG 1320  
Db 1261 GAAAGTTGTAACAAGACCGCAGAGTCAATGAACATGTGAGAGAGAGAGAGAGAGAGAG 1320  
QY 1321 CCGGAGAGAGATTCAATGACAGCATCCCAAGAACTTTGGCCTGATCGCATCTTCTGGAAG 1380  
Db 1321 CCGGAGAGAGATTCAATGACAGCATCCCAAGAACTTTGGCCTGATCGCATCTTCTGGAAG 1380  
QY 1381 GAAAGCAAGTGGCTGAGTGCGTCTCTATTACTGACTAAGAAAGATGAAACTATTA 1440

Db 1381 GAAAGCAAGTGGCTGAGTGCGTCTCTATTACTGACTAAGAAAGATGAAACTATTA 1440  
QY 1441 GAGCCTGTGAGACCGAGACTATCGGCGCGGCAAGAGCCAGACAGCAACACAGACGA 1500  
Db 1441 GAGCCTGTGAGACCGAGACTATCGGCGCGGCAAGAGCCAGACAGCAACACAGACGA 1500  
QY 1501 GCAGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1560  
Db 1501 GCAGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1560  
QY 1561 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
Db 1561 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
QY 1621 CGACAAAGAGAGAGCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
Db 1621 CGACAAAGAGAGAGCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
QY 1681 GAAAGAGCTGTGGCTTCGCAAGAGCGCGCAAACTGCAACAGCCAGGAGAGCGCAAGG 1740  
Db 1681 GAAAGAGCTGTGGCTTCGCAAGAGCGCGCAAACTGCAACAGCCAGGAGAGCGCAAGG 1740  
QY 1741 CCGCATCAACCGGCTCAATGAGCTTAATGAGCCCAACAGCGAGAGAGCCATCAACCC 1800  
Db 1741 CCGCATCAACCGGCTCAATGAGCTTAATGAGCCCAACAGCGAGAGAGCCATCAACCC 1800  
QY 1801 GAGCGCGAGCTGGCTCCATGAGAGCTGAATGAGAGTTCTGCTGGAAGAGAGAGAGAT 1860  
Db 1801 GAGCGCGAGCTGGCTCCATGAGAGCTGAATGAGAGTTCTGCTGGAAGAGAGAGAGAT 1860  
QY 1861 GGAAGCAGCCAAAGAAAGTCTCTTGGAACACGCGCGCAACTGTGTGSCCATGCGGAT 1920  
Db 1861 GGAAGCAGCCAAAGAAAGTCTCTTGGAACACGCGCGCAACTGTGTGSCCATGCGGAT 1920  
QY 1921 GGTGGGCTTCAAGACTGTGTGCAAGTGAAGACTTTCACTTCACTCAAGAAAGAGCA 1980  
Db 1921 GGTGGGCTTCAAGACTGTGTGCAAGTGAAGACTTTCACTTCACTCAAGAAAGAGCA 1980  
QY 1981 GAACTCTGATGAGATCTTGCAGACAGCAAGCTGAAGATGAGAGAGAGAGAGAGAG 2040  
Db 1981 GAACTCTGATGAGATCTTGCAGACAGCAAGCTGAAGATGAGAGAGAGAGAGAGAG 2040  
QY 2041 GAGGAAGAAAGAAAGAGCGCGCGCGCGCGCAAGAGAGAGCTGCAATTCGCGCGTGGT 2100  
Db 2041 GAGGAAGAAAGAAAGAGCGCGCGCGCGCGCGCGCAAGAGAGAGCTGCAATTCGCGCGTGGT 2100  
QY 2101 GAGGATGAGAGAGATGAGAGCGGTGCGGCGGTGAGCGGAATGAGAGAGAGATGAGAG 2160  
Db 2101 GAGGATGAGAGAGATGAGAGCGGTGCGGCGGTGAGCGGAATGAGAGAGAGATGAGAG 2160  
QY 2161 GCGTGAAGCTTTAATGCTCTGAGGATGAGGTCGCAAGAGAGAGATGAGAGAGAGAG 2220  
Db 2161 GCGTGAAGCTTTAATGCTCTGAGGATGAGGTCGCAAGAGAGAGATGAGAGAGAGAG 2220  
QY 2221 CACTGTCAACAAAGCTCAGACCCGAGAGCATCCCTCTCTCTCACTGAGAGCGGCA 2280  
Db 2221 CACTGTCAACAAAGCTCAGACCCGAGAGCATCCCTCTCTCTCACTGAGAGCGGCA 2280  
QY 2281 GGAACAGAGAGAGATGAGGCGCCAGAGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
Db 2281 GGAACAGAGAGAGATGAGGCGCCAGAGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
QY 2341 AGGCGCACCGACCCGACGAGAGAGATCCCGGCGCCGCAATTGAGAGCAACCGCGGCTC 2400  
Db 2341 AGGCGCACCGACCCGACGAGAGAGATCCCGGCGCCGCAATTGAGAGCAACCGCGGCTC 2400  
QY 2401 TGAAGCCACCGAGAGCGCTTACGCGCCCAACAGCAACCCCATGCGCTTGCACCTTCC 2460  
Db 2401 TGAAGCCACCGAGAGCGCTTACGCGCCCAACAGCAACCCCATGCGCTTGCACCTTCC 2460  
QY 2461 TTGTGTCCTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520



QY 4681 GAACATGCGGAGACCACGCGCGCTGCGAGAGGAGACCTTTTGTCTCAGCAAGCATC 4740  
DB 4681 GACCATGCGGGAGCCACGCGCGCTGCGAGAGGAGACCTTTTGTCTCAGCAAGCATC 4740  
QY 4741 CCAAGACCGGAAAGCTGAGCGCTCTGTGAGATTCGCAAGTCCCGCACAAGACCGT 4800  
DB 4741 CCAAGACCGGAAAGCTGAGCGCTCTGTGAGATTCGCAAGTCCCGCACAAGACCGT 4800  
QY 4801 GCCCGAGACACCCACACCCCATCTCGCCCTATAGACACGCTTTCGGGGGTGAGTGG 4860  
DB 4801 GCCCGAGACACCCACACCCCATCTCGCCCTATAGACACGCTTTCGGGGGTGAGTGG 4860  
QY 4861 CGTGAGACTGTATCGCAGCCACATCCCTTGCGCTTGAACCCACCTTCATACCCCGGG 4920  
DB 4861 CGTGAGACTGTATCGCAGCCACATCCCTTGCGCTTGAACCCACCTTCATACCCCGGG 4920  
QY 4921 CATCCCTCTGAGACGACGCGCTGCTTACTACTACTGCCCCGACACTGGCCCCCAACCCAC 4980  
DB 4921 CATCCCTCTGAGACGACGCGCTGCTTACTACTACTGCCCCGACACTGGCCCCCAACCCAC 4980  
QY 4981 CTACCCGACCGCTATCCACCCCTACCTCATCCGCGGCTACCCCGACACGCGCGGCTGGA 5040  
DB 4981 CTACCCGACCGCTATCCACCCCTACCTCATCCGCGGCTACCCCGACACGCGCGGCTGGA 5040  
QY 5041 GAACCGGAGACCATCATCAATGACTACATCACTTCGACAGAGATGCAACAACGCGC 5100  
DB 5041 GAACCGGAGACCATCATCAATGACTACATCACTTCGACAGAGATGCAACAACGCGC 5100  
QY 5101 CACCGCCATTTGGCCCGACGAGCTGATATGTGAGGGGGCTCTTGCCCCCGAGATCTCTGCT 5160  
DB 5101 CACCGCCATTTGGCCCGACGAGCTGATATGTGAGGGGGCTCTTGCCCCCGAGATCTCTGCT 5160  
QY 5161 GGCACTCAACTACGCTGCGGGGTCCCGAGAGCATCATCGACTGTCCCAAGTSCACACT 5220  
DB 5161 GGCACTCAACTACGCTGCGGGGTCCCGAGAGCATCATCGACTGTCCCAAGTSCACACT 5220  
QY 5221 GCTGTGCTGCTGCCCCCGACACAGGACACCCAGACCAACCGCATGACACCGCTTGCTTA 5280  
DB 5221 GCTGTGCTGCTGCCCCCGACACAGGACACCCAGACCAACCGCATGACACCGCTTGCTTA 5280  
QY 5281 CCTCCCGACCGGCGCCCAAGCCCTTACAGACGCGCCACAGAGCTTCCCATCTTCCCCAGG 5340  
DB 5281 CCTCCCGACCGGCGCCCAAGCCCTTACAGACGCGCCACAGAGCTTCCCATCTTCCCCAGG 5340  
QY 5341 AGGTCCAAACAACCTTGACAAACCAACCAACGCTCTGCTCGAGCGGGAGCGAGACCG 5400  
DB 5341 AGGTCCAAACAACCTTGACAAACCAACCAACGCTCTGCTCGAGCGGGAGCGAGACCG 5400  
QY 5401 GGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAAGTCCATCTCAAGTCCACAGAC 5460  
DB 5401 GGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAAGTCCATCTCAAGTCCACAGAC 5460  
QY 5461 GGTGAGACACCGCATCTTGAAGACTGTGTACAGAGACAGACGGGAGAGAGGGGAG 5520  
DB 5461 GGTGAGACACCGCATCTTGAAGACTGTGTACAGAGACAGACGGGAGAGAGGGGAG 5520  
QY 5521 CAGCGCGGGGGGTGGGGGACAGACAGCGCGCGCTTCCCACTCCATGCCCCACAGAC 5580  
DB 5521 CAGCGCGGGGGGTGGGGGACAGACAGCGCGCGCTTCCCACTCCATGCCCCACAGAC 5580  
QY 5581 CTGCGCCATCTTCCCTCGGACCCAGAGATGCTCTTCAGACAGAGACCGAGTGTTCACAA 5640  
DB 5581 CTGCGCCATCTTCCCTCGGACCCAGAGATGCTCTTCAGACAGAGACCGAGTGTTCACAA 5640  
QY 5641 CACAGGCAATGAAGGGTATCATACCGCTGTGAGAGCCCAAGCCACAGCTCTAGGTC 5700  
DB 5641 CACAGGCAATGAAGGGTATCATACCGCTGTGAGAGCCCAAGCCACAGCTCTAGGTC 5700  
QY 5701 CACCTCAACTCTCTACCGGCTTCCGCGAGCTGCAATTCCTGACACCCCACTGCGCC 5760  
DB 5701 CACCTCAACTCTCTACCGGCTTCCGCGAGCTGCAATTCCTGACACCCCACTGCGCC 5760  
QY 5761 ACTGGGCGGACCCCTGATGGGGGTCTACCTTACCTCATGAGCCCGTCTTGCTGCCAA 5820

DB 5761 ACTGGGCGGACCCCTGATGGGGGTCTACCTTACCTCATGAGCCCGTCTTGCTGCCAA 5820  
QY 5821 GGAAGCCCCCGGGGTGCGCCGGCCAGAGCGGCGCCGACAGACACCGGACATGCTTCTCT 5880  
DB 5821 GGAAGCCCCCGGGGTGCGCCGGCCAGAGCGGCGCCGACAGACACCGGACATGCTTCTCT 5880  
QY 5881 CGCCAAAGCCCCAGCGCGCTCGGGCTGGAAGCCGCGCTCTTCCCGACAAAGGCTCGGA 5940  
DB 5881 CGCCAAAGCCCCAGCGCGCTCGGGCTGGAAGCCGCGCTCTTCCCGACAAAGGCTCGGA 5940  
QY 5941 GCCCGGCGCTTATGCTCTCTGTCTCTGCGCACGACCATTCGCGGACACCCCTGCGAA 6000  
DB 5941 GCCCGGCGCTTATGCTCTCTGTCTCTGCGCACGACCATTCGCGGACACCCCTGCGAA 6000  
QY 6001 GAACCTGCACTTACACAGCGCAGCGCGGACCGCGGCGCACCTGCTGCGCTCGGA 6060  
DB 6001 GAACCTGCACTTACACAGCGCAGCGCGGACCGCGGCGCACCTGCTGCGCTCGGA 6060  
QY 6061 CCGGACCGGGAAAAAGACTCAAAAGTAAACCTTTTTCATCGAGAACTGAACTCCGTTT 6120  
DB 6061 CCGGACCGGGAAAAAGACTCAAAAGTAAACCTTTTTCATCGAGAACTGAACTCCGTTT 6120  
QY 6121 TCTGGGTTACCAAGCGCAGACGCTACAGCTCCGAAAGGGGTGAGGCTCTGACGCTGTGAG 6180  
DB 6121 TCTGGGTTACCAAGCGCAGACGCTACAGCTCCGAAAGGGGTGAGGCTCTGACGCTGTGAG 6180  
QY 6181 CTACCCGAGTGCACCCACAGCAAGGGGCTCCCAAGACCTGGAAGAGCTGCACAAAG 6240  
DB 6181 CTACCCGAGTGCACCCACAGCAAGGGGCTCCCAAGACCTGGAAGAGCTGCACAAAG 6240  
QY 6241 CCACCTGAGGGGGAGGCTGCGGCGCCCAAGCAGCGCCGCTGGAAGCTTGGCGGGAGGC 6300  
DB 6241 CCACCTGAGGGGGAGGCTGCGGCGCCCAAGCAGCGCCGCTGGAAGCTTGGCGGGAGGC 6300  
QY 6301 CGCCCACTTCCCAACCTTGGCGCGCTGCTGAGAGCCAGCCCTGTGTCAGCCCGCTGCT 6360  
DB 6301 CGCCCACTTCCCAACCTTGGCGCGCTGCTGAGAGCCAGCCCTGTGTCAGCCCGCTGCT 6360  
QY 6361 CGAGCGCGCCCGAGGGGTCAAAAGGTACACAGCGGGTGTACCCCTGGGCGACACATGAG 6420  
DB 6361 CGAGCGCGCCCGAGGGGTCAAAAGGTACACAGCGGGTGTACCCCTGGGCGACACATGAG 6420  
QY 6421 TGAAGTCACTACACAGGACTACACCCGACACACCCACAGAGCTCAGCGACCCCTGCGC 6480  
DB 6421 TGAAGTCACTACACAGGACTACACCCGACACACCCACAGAGCTCAGCGACCCCTGCGC 6480  
QY 6481 CGCCCGCTTCTACTCTTCCCTGGGGCCAGCTGCGCCGCTCTGAGACTTCGCGCGCCACG 6540  
DB 6481 CGCCCGCTTCTACTCTTCCCTGGGGCCAGCTGCGCCGCTCTGAGACTTCGCGCGCCACG 6540  
QY 6541 CAGTGAACCTTACCTCCGCGCCCGGACCATGTGACCCCGGCGGTGCTCCCCACAG 6600  
DB 6541 CAGTGAACCTTACCTCCGCGCCCGGACCATGTGACCCCGGCGGTGCTCCCCACAG 6600  
QY 6601 CGAAGGGGGCAAGAGTCTCGAGAGCCAAACAAGCTGTGGTCTTGGTGTGTGAGGA 6660  
DB 6601 CGAAGGGGGCAAGAGTCTCGAGAGCCAAACAAGCTGTGGTCTTGGTGTGTGAGGA 6660  
QY 6661 CGGTATTGAACCTGTGTCTCCACCGGAGGGCATGACGAGCCAGGGCACTCCCGAGTGC 6720  
DB 6661 CGGTATTGAACCTGTGTCTCCACCGGAGGGCATGACGAGCCAGGGCACTCCCGAGTGC 6720  
QY 6721 TGTGTACCCGCTGCTGATCCGGAGATGGGGAAACAGCGAGGCCAGAGATGGGCTCCAA 6780  
DB 6721 TGTGTACCCGCTGCTGATCCGGAGATGGGGAAACAGCGAGGCCAGAGATGGGCTCCAA 6780  
QY 6781 GTCTCAGGCAACACAGCGCGCGCACTTCTTCAAGAGCTGACCGAGAGCAATC 6840  
DB 6781 GTCTCAGGCAACACAGCGCGCGCACTTCTTCAAGAGCTGACCGAGAGCAATC 6840  
QY 6841 CGCCATGCTCAAGTCCAAAGAAAGAGATCAAGAACTGAAACCCACACCGGAA 6900



6841 CGCCATGCTCAAGTCCAGAGAGAGATCAACAAGAGTGAACCCCAACCGGAA 6900  
 6901 TGAGCTGAATACATATACAGAGCTGGGAGCGAGATCTTCAATATGCGGCATCAC 6960  
 6901 TGAGCTGAATACATATACAGAGCTGGGAGCGAGATCTTCAATATGCGGCATCAC 6960  
 6961 CGGAACAGGCTTATATACCTATAGAAAGCAGGCGGTGCAGGAACATGCGACCAACAT 7020  
 6961 CGGAACAGGCTTATATACCTATAGAAAGCAGGCGGTGCAGGAACATGCGACCAACAT 7020  
 7021 GGGGCTGAGAGGCTTATATAGAAAGCATCATGGTAAATATGACCAAGTGGAGAGTCT 7080  
 7021 GGGGCTGAGAGGCTTATATAGAAAGCATCATGGTAAATATGACCAAGTGGAGAGTCT 7080  
 7081 CCGGCGCTCAGGCGCAATGCTTTAACTCTGAATGCGAGGCGAGGCTGCGGCTGC 7140  
 7081 CCGGCGCTCAGGCGCAATGCTTTAACTCTGAATGCGAGGCGAGGCTGCGGCTGC 7140  
 7141 TATGCCATTAACCGCTGTCAGCGAGTGAACAACATCTGACCTGCGCAGGTGGCG 7200  
 7141 TATGCCATTAACCGCTGTCAGCGAGTGAACAACATCTGACCTGCGCAGGTGGCG 7200  
 7201 CGGGAAGGCGCAAGTCTCTGAGAGACCCAGACCGGAAAGCAAGTCCCGGCGCGG 7260  
 7201 CGGGAAGGCGCAAGTCTCTGAGAGACCCAGACCGGAAAGCAAGTCCCGGCGCGG 7260  
 7261 CCGGAGTCTGGGAGCGGCGACCCCTCTGCTCTGATGCACTCGGAGGGAACCTGCA 7320  
 7261 CCGGAGTCTGGGAGCGGCGACCCCTCTGCTCTGATGCACTCGGAGGGAACCTGCA 7320  
 7321 CCGGCGGAGCGGCTGTCACCAACCGGCTGTCAGAGAGACAGGCTCTGCGCAGGTTCAC 7380  
 7321 CCGGCGGAGCGGCTGTCACCAACCGGCTGTCAGAGAGACAGGCTCTGCGCAGGTTCAC 7380  
 7381 GGCATTCCTCTCAACACCCCTGATCATGCGGCTGTCAGAGCGGAGTGTATGCTTCCAC 7440  
 7381 GGCATTCCTCTCAACACCCCTGATCATGCGGCTGTCAGAGCGGAGTGTATGCTTCCAC 7440  
 7441 CCGACCGGCGCTCCCGCGGAGCGGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 7500  
 7441 CCGACCGGCGCTCCCGCGGAGCGGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 7500  
 7501 GGAGCCCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7560  
 7501 GGAGCCCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7560  
 7561 AACAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7620  
 7561 AACAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7620  
 7621 CCGGCTTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7680  
 7621 CCGGCTTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7680  
 7611 CCGGCTTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7670  
 7611 CCGGCTTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7670  
 7681 CCGGCTTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7740  
 7681 CCGGCTTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7740  
 7671 CCGGCTTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7730  
 7671 CCGGCTTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7730  
 7741 AACTAAGATCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7800  
 7741 AACTAAGATCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7800  
 7731 AACTAAGATCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7790  
 7731 AACTAAGATCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7790  
 7801 GTCCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7860  
 7801 GTCCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7860  
 7791 GTCCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7850  
 7791 GTCCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7850  
 7861 ACAAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7920  
 7861 ACAAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7920  
 7851 ACAAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7910  
 7851 ACAAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7910  
 7921 TGCCTTCCCGCAATGATTTGGAACCAAGTCTAACTGAGGCTGCGAGGCGGCGGCGCTT 7980  
 7921 TGCCTTCCCGCAATGATTTGGAACCAAGTCTAACTGAGGCTGCGAGGCGGCGGCGCTT 7980  
 7911 TGCCTTCCCGCAATGATTTGGAACCAAGTCTAACTGAGGCTGCGAGGCGGCGGCGCTT 7970  
 7911 TGCCTTCCCGCAATGATTTGGAACCAAGTCTAACTGAGGCTGCGAGGCGGCGGCGCTT 7970

7981 CCCTCCGCTCCCATCCGCTTATAGCGCTGACAGATGAGCAGAGGCTGTCCAGGCC 8040  
 7971 CCCTCCGCTCCCATCCGCTTATAGCGCTGACAGATGAGCAGAGGCTGTCCAGGCC 8030  
 8041 CCAAGTGGCTGCTCCGCTGCTCCCAAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 8100  
 8031 CCAAGTGGCTGCTCCGCTGCTCCCAAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 8090  
 8101 GTCAGGCGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8160  
 8091 GTCAGGCGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8150  
 8161 GGAAGTGAATGCTTTTTCACACATGCTGCGCAGCGGTCGAGGAGGAGGAGGAGGAGGAG 8220  
 8151 GGAAGTGAATGCTTTTTCACACATGCTGCGCAGCGGTCGAGGAGGAGGAGGAGGAGGAG 8210  
 8221 AATGATGTGTGTGTTTACAGGGTAAATTTTGTATATCTTCAATGAATTAATGAGTGT 8280  
 8211 AATGATGTGTGTGTTTACAGGGTAAATTTTGTATATCTTCAATGAATTAATGAGTGT 8270  
 8281 TTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8340  
 8271 TTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8330  
 8341 TTCAGAGGCGGTGTCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8400  
 8331 TTCAGAGGCGGTGTCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8390  
 8401 GGGAGCTGCTGTCAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8460  
 8391 GGGAGCTGCTGTCAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8450  
 8461 TTCAGTGTGATTTCTGTCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8520  
 8451 TTCAGTGTGATTTCTGTCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8510  
 8521 CGTGTCTTAATTAATAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8571  
 8511 CGTGTCTTAATTAATAAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8561

RESULT 2  
 US-11-052-344-1  
 ; Sequence 1, Application US/11052344  
 ; Publication No. US20050191674A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Colimet, Lionel  
 ; TITLE OF INVENTION: Method Of Prognosis Of Metastasis By Detection Of FRA12E Fragile  
 ; TITLE OF INVENTION: Method Of Prognosis Of Metastasis By Detection Of FRA12E Fragile  
 ; FILE REFERENCE: 03551.0181  
 ; CURRENT APPLICATION NUMBER: US/11/052,344  
 ; PRIOR FILING DATE: 2005-02-07  
 ; PRIOR FILING DATE: 2004-02-06  
 ; NUMBER OF SEQ ID NOS: 1  
 ; SEQ ID NO 1 8561  
 ; LENGTH: 8561  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: The open reading frame starts at nucleotide 2 and ends at nucleot  
 ; US-11-052-344-1

Query Match 99.7%; Score 8541; DB 10; Length 8561;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 8560; Conservative 1; Mismatches 0; Indels 10; Gaps 1;  
 1 CATGTGCGGCTCCACACAGCTTGTGCAACAGCTGAGAGGCGCATGAGCCCGCTACCC 60  
 1 CATGTGCGGCTCCACACAGCTTGTGCAACAGCTGAGAGGCGCATGAGCCCGCTACCC 60

QY 61 GCCCAGACGCTTTCTTACCCAGTCAGATCGCCCGAGCGCAACGAGCTGTGGCTCCT 120  
Db 61 GCCCAGACGCTTTCTTACCCAGTCAGATCGCCCGAGCGCAACGAGCTGTGGCTCCT 120  
QY 121 GAGAGTCCAGACGCACTCCCGGACTATATGCTCCCACTGTGTGGGGGTCCATATCCA 180  
Db 121 GAGAGTCCAGACGCACTCCCGGACTATATGCTCCCACTGTGTGGGGGTCCATATCCA 180  
QY 181 GCCCCAGCGGCGGAGGCGCTCCCTGCTGTCTGAGTTCCAGCCCGGAAATGAACGGTCCCA 240  
Db 181 GCCCCAGCGGCGGAGGCGCTCCCTGCTGTCTGAGTTCCAGCCCGGAAATGAACGGTCCCA 240  
QY 241 GAGAGTCCACCTGTGCGGCGCAAGTCCCATCTCATATCTGTCCCGAGCTGTGGGAACTCAGAGAT 300  
Db 241 GAGAGTCCACCTGTGCGGCGCAAGTCCCATCTCATATCTGTCCCGAGCTGTGGGAACTCAGAGAT 300  
QY 301 GAGATTCAATTGAAAGCAGACGCGCTTCGGCTTAGAGCTGTGCTGTACCCCTGTCTGTGAC 360  
Db 301 GAGATTCAATTGAAAGCAGACGCGCTTCGGCTTAGAGCTGTGCTGTGTGACCCCTGTCTGTGAC 360  
QY 361 GTCAACCCCTGTGGGCGACGGGGCGAGCTGTGCGGGATCTGAAAGACTCAACCAAGACCGTAG 420  
Db 361 GTCAACCCCTGTGGGCGACGGGGCGAGCTGTGCGGGATCTGAAAGACTCAACCAAGACCGTAG 420  
QY 421 CCTGACGGGCGAAGCTGGAACCGGTGTCTTCCCGCAGCCCGCCGCAACTGAACCTGAAGCT 480  
Db 421 CCTGACGGGCGAAGCTGGAACCGGTGTCTTCCCGCAGCCCGCCGCAACTGAACCTGAAGCT 480  
QY 481 GAGAGTGGTGGCCGCAACGGCTGTCCAGAGAGAGCTGATCCGAACCTGTGACCCGGTGGGA 540  
Db 481 GAGAGTGGTGGCCGCAACGGCTGTCCAGAGAGAGCTGATCCGAACCTGTGACCCGGTGGGA 540  
QY 541 CCGAGAGATCAACATGTGTAGAGCAGACAGATCTCTAAGCTGAAGAAAGAACAGCAAGCT 600  
Db 541 CCGAGAGATCAACATGTGTAGAGCAGACAGATCTCTAAGCTGAAGAAAGAACAGCAAGCT 600  
QY 601 GAGAGAGAGAGGCTGTGCCAAGCCCGCCGAGCTGTGAGAACCCCGTGTCAACCGCCCATGGA 660  
Db 601 GAGAGAGAGAGGCTGTGCCAAGCCCGCCGAGCTGTGAGAACCCCGTGTCAACCGCCCATGGA 660  
QY 661 GTCGAAGCACCAGACCTGTGTGCAAGTCAATCGACGAAACCCGAGAGAGGCTGGAAGC 720  
Db 661 GTCGAAGCACCAGACCTGTGTGCAAGTCAATCGACGAAACCCGAGAGAGGCTGGAAGC 720  
QY 721 TGCAATGCGGATTTCTGGAAGGCTGTGGGGCCCGAGGTGAGCTGTGCGCTGTACAACAGGC 780  
Db 721 TGCAATGCGGATTTCTGGAAGGCTGTGGGGCCCGAGGTGAGCTGTGCGCTGTACAACAGGC 780  
QY 781 CTCCGACACCCCGGAGATCATGAGAACATCAAAATAAACAGGCGATGCGGAGAGAGCT 840  
Db 781 CTCCGACACCCCGGAGATCATGAGAACATCAAAATAAACAGGCGATGCGGAGAGAGCT 840  
QY 841 AATCTTTGTAATTCAAGAGAGAGATCAAGCTGTGGAACATGGAAGCAGAGATTCTTGCCA 900  
Db 841 AATCTTTGTAATTCAAGAGAGAGATCAAGCTGTGGAACATGGAAGCAGAGATTCTTGCCA 900  
QY 901 GGGGCTATGACCAAGCTCATGGAAGGCTTTGGAAAAAAAGGTGAGCGGATGAAAAACAACC 960  
Db 901 GGGGCTATGACCAAGCTCATGGAAGGCTTTGGAAAAAAAGGTGAGCGGATGAAAAACAACC 960  
QY 961 GGGCGCGCGGCGCAAGAGAGCAAGTGTGCGCAGATCTACGAAAAAGCAATTCCTGAGAT 1020  
Db 961 GGGCGCGCGGCGCAAGAGAGCAAGTGTGCGCAGATCTACGAAAAAGCAATTCCTGAGAT 1020  
QY 1021 CCGCAAGCAGCGCAGCTGTGAGAGCGCATGTGCAAGCAGAGGTTGGCCAGCGGCGCGAGTGG 1080  
Db 1021 CCGCAAGCAGCGCAGCTGTGAGAGCGCATGTGCAAGCAGAGGTTGGCCAGCGGCGCGAGTGG 1080  
QY 1081 GGTGTTCATGTGGCGCGCCCGAGAGAGCAAGAGGTGTCAAGAGATCTGTGAGGCTCTTC 1140  
Db 1081 GGTGTTCATGTGGCGCGCCCGAGAGAGCAAGAGGTGTCAAGAGATCTGTGAGGCTCTTC 1140  
QY 1141 AGAGCAGAGAACTGTAGAGAAAGATGTGCGCAGCTGTGATCCCGCATGCTGTA 1200

Db 1141 AGAGCAGAGAACTGTAGAGAAAGATGTGCGCAGCTGTGATCCCGCATGCTGTA 1200  
QY 1201 CGAGCTGACCAAGCAGCGCATCAAGTTTCATATCAATGAACGGGCTTATGTCCGACCCAT 1260  
Db 1201 CGAGCTGACCAAGCAGCGCATCAAGTTTCATATCAATGAACGGGCTTATGTCCGACCCAT 1260  
QY 1261 GAAAGTGTACAAGACCGCCAGTTCATGACATGTGTGAGTGTGACAGAGAAAGAGACCTT 1320  
Db 1261 GAAAGTGTACAAGACCGCCAGTTCATGACATGTGTGAGTGTGACAGAGAAAGAGACCTT 1320  
QY 1321 CCGGAGAAAGTTTCATGCAACATCCCAAGAACTTTGGCTGATCGCATATTTCTTGAAGG 1380  
Db 1321 CCGGAGAAAGTTTCATGCAACATCCCAAGAACTTTGGCTGATCGCATATTTCTTGAAGG 1380  
QY 1381 GAAAGCAGTGGCTGATGTGTGCTCTATTTAATCACTGTAAGAAAGATGAACTTATA 1440  
Db 1381 GAAAGCAGTGGCTGATGTGTGCTCTATTTAATCACTGTAAGAAAGATGAACTTATA 1440  
QY 1441 GAGCTGTGTGAGACGGAAGCTATGTGCGCGCGGCGCAAGGCGAGCAACACAGACGA 1500  
Db 1441 GAGCTGTGTGAGACGGAAGCTATGTGCGCGCGGCGCAAGGCGAGCAACACAGACGA 1500  
QY 1501 GCAGCAGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1560  
Db 1501 GCAGCAGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1560  
QY 1561 AGATGAGAAAGAGAAAGAAAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
Db 1561 AGATGAGAAAGAGAAAGAAAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
QY 1621 GCACAGAGAGACCTCTCAAGAGAGAGACAGACGACCTCAAGGAGAGAGACGACGAG 1680  
Db 1621 GCACAGAGAGACCTCTCAAGAGAGAGAGACGACGAGCTCAAGGAGAGAGACGACGAG 1680  
QY 1681 GAAAGAGGCTGTGTCTCAAAAGGCGCAAACTGTCCAAACAGCCAGGAGAGAGCGAAAGG 1740  
Db 1681 GAAAGAGGCTGTGTGTCTCAAAAGGCGCAAACTGTCCAAACAGCCAGGAGAGAGCGAAAGG 1740  
QY 1741 CCGCATCAACCGGCTCAATGTGCTAATGAGGCCCAAGGAGGAGGCTATCAACCCCGACGA 1800  
Db 1741 CCGCATCAACCGGCTCAATGTGCTAATGAGGCCCAAGGAGGAGGCTATCAACCCCGACGA 1800  
QY 1801 GAGCGCCGAGCTGTGCTCATGAGCTGATGAGAGTCTGTGCTGGAACGAAAGAAAT 1860  
Db 1801 GAGCGCCGAGCTGTGCTCATGAGCTGATGAGAGTCTGTGCTGGAACGAAAGAAAT 1860  
QY 1861 GGAACACGCCAAGAAAGGTCTCTGTGAACAGCGCGCAACTGTGTGCGCCATGCGCCGAT 1920  
Db 1861 GGAACACGCCAAGAAAGGTCTCTGTGAACAGCGCGCAACTGTGTGCGCCATGCGCCGAT 1920  
QY 1921 GGTGGGCTCCAAAGCTGTGTGCGCAGTGTGAAGCTTCACTTCAACTCAAGAAAGGCA 1980  
Db 1921 GGTGGGCTCCAAAGCTGTGTGCGCAGTGTGAAGCTTCACTTCAACTCAAGAAAGGCA 1980  
QY 1981 GAACTCTGATGAGATCTTTCAGCAGCACAAGCTGAAATGTGAGAGAGAGAGAGAGAGAG 2040  
Db 1981 GAACTCTGATGAGATCTTTCAGCAGCACAAGCTGAAATGTGAGAGAGAGAGAGAGAGAG 2040  
QY 2041 GAGGAAAGAAAGAAAGCGCGCGCGCGCGCAAGGAGAGGCTGCAATTCGCGCCGTGCT 2100  
Db 2041 GAGGAAAGAAAGAAAGCGCGCGCGCGCGCGCAAGGAGAGGCTGCAATTCGCGCCGTGCT 2100  
QY 2101 GAGAGATGAGAGAGATGAGAGGCTGTGCGGCGGTGACGCGGAAATGAGAGATGTGTGAGGA 2160  
Db 2101 GAGAGATGAGAGAGATGAGAGGCTGTGCGGCGGTGACGCGGAAATGAGAGATGTGTGAGGA 2160  
QY 2161 GAGTGAAGCTTTAATGCTCTGTGGAATGAGGTGCCAGAGGAGAAATGAGAGGCGCCAGC 2220  
Db 2161 GAGTGAAGCTTTAATGCTCTGTGGAATGAGGTGCCAGAGGAGAAATGAGAGGCGCCAGC 2220  
QY 2221 CACTGTCAACAAACAGCTGACACCGAGAGCATTCCTCTCTCACTGAGGCGCCAA 2280

[illegible]

QY	3361	GCAAAATAGTGTSCCATCTCTCCCAAGGAATGTGTGTCTCAAGTCTCAAGTCTCCCTGTACTCAAGCA	3420
Db	3361	GCAATATGTGTSCCATCTCTCCCAAGGAATGTGTGTCTCAAGTCTCAAGTCTCCCTGTACTCAAGCA	3420
QY	3421	TGCGAAGGCCCCGGGTGGGGCCCTGTCACTATGGGGCTGCCCTGGCCATTTGGACCCCAAAA	3480
Db	3421	TGCGAAGGCCCCGGGTGGGGCCCTGTCACTATGGGGCTGCCCTGGCCATTTGGACCCCAAAA	3480
QY	3481	GCTGGCAACCCCTTCAGCCGAGTGAAGCAGAGCAGACTGTCCCAAGGGGACAGGCTGTGGGC	3540
Db	3481	GCTGGCAACCCCTTCAGCCGAGTGAAGCAGAGCAGACTGTCCCAAGGGGACAGGCTGTGGGC	3540
QY	3541	ACCGAAGAGCTGTGGGGGTGCCCAAGCCAGAGAGGCTCGTGTCTGAAGAGGACAAGCTCT	3600
Db	3541	ACCGAAGAGCTGTGGGGGTGCCCAAGCCAGAGAGGCTCGTGTCTGAAGAGGACAAGCTCT	3600
QY	3601	GGGCTCAAGTTCGGGGCCGAAAGCATCAACAAAGCATTCGACGACAAGGGGTGCCCTGGGA	3660
Db	3601	GGGCTCAAGTTCGGGGCCGAAAGCATCAACAAAGCATTCGACGACAAGGGGTGCCCTGGGA	3660
QY	3661	CAGGCGCATCATATACCGCGGGCTCATATCAACGACGACGACAGCTAGCTGCTGTCAAA	3720
Db	3661	CAGGCGCATCATATACCGCGGGCTCATATCAACGACGACGACAGCTAGCTGCTGTCAAA	3720
QY	3721	GGGCAACATCAACAGATTCAGCTCGGCGAGGACAGCCGACAGTCTGTGGACCGCGGCCGGA	3780
Db	3721	GGGCAACATCAACAGATTCAGCTCGGCGAGGACAGCCGACAGTCTGTGGACCGCGGCCGGA	3780
QY	3781	GGACAGCTCTGCCCAAGGACGACGTCACTCTACGAAGGACAAAGAGGGCCACGTCTTGTCTTA	3840
Db	3781	GGACAGCTCTGCCCAAGGACGACGTCACTCTACGAAGGACAAAGAGGGCCACGTCTTGTCTTA	3840
QY	3841	TGAAGGTGTGCATGTCTGTGAACCCAGTGTCTCCAAAGAGAGAGCGGCAAGAGCTCAGAAC	3900
Db	3841	TGAAGGTGTGCATGTCTGTGAACCCAGTGTCTCCAAAGAGAGAGCGGCAAGAGCTCAGAAC	3900
QY	3901	CCCCCATGAGACGGCCCGCCCCCAAGGACGACCTTATGATGATGAGAGGGCCGCTGTGGCAG	3960
Db	3901	CCCCCATGAGACGGCCCGCCCCCAAGGACGACCTTATGATGATGAGAGGGCCGCTGTGGCAG	3960
QY	3961	AGCATCTCTCTCAGCCAGCATCGAAGGTCTCATGGGCGGTGCCATCCCGCCGGAGCGACA	4020
Db	3961	AGCATCTCTCTCAGCCAGCATCGAAGGTCTCATGGGCGGTGCCATCCCGCCGGAGCGACA	4020
QY	4021	CAGCCCCCAACCACTCTCAAAAGAGCAGCAACATCCGGGGGTCCATCAACAAGGAAATCCC	4080
Db	4021	CAGCCCCCAACCACTCTCAAAAGAGCAGCAACATCCGGGGGTCCATCAACAAGGAAATCCC	4080
QY	4081	TGGGTCTCTAGTGAAGGACAGAGAGACTACCTGTGTGGAGGGCCAAAGTCTCTAAAGCG	4140
Db	4081	TGGGTCTCTAGTGAAGGACAGAGAGACTACCTGTGTGGAGGGCCAAAGTCTCTAAAGCG	4140
QY	4141	GGAGGGACAGCTCTCGCCCCCAACCGCCCTCAAGGACCTGACCGAGGCTTACAAGACGCA	4200
Db	4141	GGAGGGACAGCTCTCGCCCCCAACCGCCCTCAAGGACCTGACCGAGGCTTACAAGACGCA	4200
QY	4201	GGCCCTTGGGGCCCCCTGGAAGCTGGAAGCCGGGCCATGGAAGGCTGTGGTGCACAGGTGAAGGA	4260
Db	4201	GGCCCTTGGGGCCCCCTGGAAGCTGGAAGCCGGGCCATGGAAGGCTGTGGTGCACAGGTGAAGGA	4260
QY	4261	GGGGGGGCGGCTCATTCATGAGATCCCGGGGCGAGAGACTGCGGGGACACGGCCCGAGCTGCC	4320
Db	4261	GGGGGGGCGGCTCATTCATGAGATCCCGGGGCGAGAGACTGCGGGGACACGGCCCGAGCTGCC	4320
QY	4321	CTTGGCCCCGGGCGGCTCAAGAGAGGGCTCATTCAGCAGAGGGACCCCGGCTCAAGTACGA	4380
Db	4321	CTTGGCCCCGGGCGGCTCAAGAGAGGGCTCATTCAGCAGAGGGACCCCGGCTCAAGTACGA	4380
QY	4381	CACCGGCGGCTCCACCACTGTGTCTCAAAAAGCAGAGTACGCTCTCTCATTCGGCAGGCC	4440
Db	4381	CACCGGCGGCTCCACCACTGTGTCTCAAAAAGCAGAGTACGCTCTCTCATTCGGCAGGCC	4440

4441 CGGCGGAGCTTCCACCCGTCACCCGCTGATGTATGTCGCAAGCCCGGGACTGGA 4500  
4441 CGGCGGAGCTTCCACCCGTCACCCGCTGATGTATGTCGCAAGCCCGGGACTGGA 4500  
4501 ACGTGCCTGCTACGAGGAGAGGCTGAAGAGCGGGCAAGGACCGGACAGAGCTCGGGGGG 4560  
4501 ACGTGCCTGCTACGAGGAGAGGCTGAAGAGCGGGCAAGGACCGGACAGAGCTCGGGGGG 4560  
4561 CTTCATTGCGCGCGGCGCCCGGTCATTGTGCTGTAGCTGGTAAAGCCGCGCAGAGCCC 4620  
4561 CTTCATTGCGCGGCGGCGCCCGGTCATTGTGCTGTAGCTGGTAAAGCCGCGCAGAGCCC 4620  
4621 CCTGACTATGAGGACACAGGGGCAACCTTTGCGGGCACTTCCACAGAGTTGCGCGCT 4680  
4621 CCTGACTATGAGGACACAGGGGCAACCTTTGCGGGCACTTCCACAGAGTTGCGCGCT 4680  
4661 GACCAATGCGGGGAGCCACGCGCGCTGAGGAGGGGAGCCCTTTGTCAGAGAAAGGCAATC 4740  
4661 GACCAATGCGGGGAGCCACGCGCGCTGAGGAGGGGAGCCCTTTGTCAGAGAAAGGCAATC 4740  
4681 GACCAATGCGGGGAGCCACGCGCGCTGAGGAGGGGAGCCCTTTGTCAGAGAAAGGCAATC 4740  
4741 CCAGGACCGAAAGCTGACGTCGACGCTTCGTAGATTCGCAAGTCCCGCAAGACAGCGT 4800  
4741 CCAGGACCGAAAGCTGACGTCGACGCTTCGTAGATTCGCAAGTCCCGCAAGACAGCGT 4800  
4801 GCCCGAGACACACCAACCCCACTTCGCTTATGAGCACTGCTTCGGGGGCTGATGG 4860  
4801 GCCCGAGACACACCAACCCCACTTCGCTTATGAGCACTGCTTCGGGGGCTGATGG 4860  
4801 GCCCGAGACACACCAACCCCACTTCGCTTATGAGCACTGCTTCGGGGGCTGATGG 4860  
4861 GGTGAGACCTGTATGAGAGCAGATCCCGCTGGCTTTCGACCCCACTCCATACCCGCGG 4920  
4861 GGTGAGACCTGTATGAGAGCAGATCCCGCTGGCTTTCGACCCCACTCCATACCCGCGG 4920  
4861 GGTGAGACCTGTATGAGAGCAGATCCCGCTGGCTTTCGACCCCACTCCATACCCGCGG 4920  
4921 CATCCCTTGAAGAGAGCGCTGCTCTTACCTGCGCCGCAACTTGGCCCCCAACCCCAAC 4980  
4921 CATCCCTTGAAGAGAGCGCTGCTCTTACCTGCGCCGCAACTTGGCCCCCAACCCCAAC 4980  
4921 CATCCCTTGAAGAGAGCGCTGCTCTTACCTGCGCCGCAACTTGGCCCCCAACCCCAAC 4980  
4981 CTACCCGCACTGTATGCTGACCTTACCTTACCTGCGCGCTTACCCGCAACCGCGCGCTGGA 5040  
4981 CTACCCGCACTGTATGCTGACCTTACCTTACCTGCGCGCTTACCCGCAACCGCGCGCTGGA 5040  
5041 GAACGGGAGACCAATCATATGACTATGACTATGACTATGACTATGACTATGACTATGACTATG 5100  
5041 GAACGGGAGACCAATCATATGACTATGACTATGACTATGACTATGACTATGACTATGACTATG 5100  
5041 GAACGGGAGACCAATCATATGACTATGACTATGACTATGACTATGACTATGACTATGACTATG 5100  
5101 CACCGGCATGCGCCAGAGAGCTGATATGCTGAGGGGCTCTGCGCCGCGAGTCTCTCGCT 5160  
5101 CACCGGCATGCGCCAGAGAGCTGATATGCTGAGGGGCTCTGCGCCGCGAGTCTCTCGCT 5160  
5161 GGCATCTCAACTATGCTGCGGGGTCCCGAGGCAATCATGCACTGTCCCAAGTGCACACCT 5220  
5161 GGCATCTCAACTATGCTGCGGGGTCCCGAGGCAATCATGCACTGTCCCAAGTGCACACCT 5220  
5221 GCTGTGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 5280  
5221 GCTGTGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 5280  
5221 GCTGTGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 5280  
5281 CTTCCCAACCGGCGCCAGGCTTTAGAGAGCGGCAAGAGAGCTTCCCACTCTCCCAAGG 5340  
5281 CTTCCCAACCGGCGCCAGGCTTTAGAGAGCGGCAAGAGAGCTTCCCACTCTCCCAAGG 5340  
5281 CTTCCCAACCGGCGCCAGGCTTTAGAGAGCGGCAAGAGAGCTTCCCACTCTCCCAAGG 5340  
5341 AGGTCCAAACACTTGAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5400  
5341 AGGTCCAAACACTTGAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5400  
5401 GGATGAGAGCGGGAGCCGGGATTCGGGAGGGGAGAAAGTTCATCTCACTCACTCACTCACTCA 5460  
5401 GGATGAGAGCGGGAGCCGGGATTCGGGAGGGGAGAAAGTTCATCTCACTCACTCACTCACTCA 5460  
5401 GGATGAGAGCGGGAGCCGGGATTCGGGAGGGGAGAAAGTTCATCTCACTCACTCACTCACTCA 5460  
5461 GGTGAGAGCGGACCACTTGTGAGACTTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5520  
5461 GGTGAGAGCGGACCACTTGTGAGACTTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5520  
5461 GGTGAGAGCGGACCACTTGTGAGACTTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5520  
5521 CAGCGCGGGGGGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5580

5521 CAGCGCGGGGGTGGGGGACAGAGAGCGCGCGCTCCACTCCATGACCAACAGCA 5580  
5581 CTGCGGCATCTTCCCTCGGAGCCAGAGATGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5640  
5581 CTGCGGCATCTTCCCTCGGAGCCAGAGATGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5640  
5641 CACAGGCATGAAGGATATCATACCGCTGTGAGGCCAGAGCCCAAGGCTTCGAGGTC 5700  
5641 CACAGGCATGAAGGATATCATACCGCTGTGAGGCCAGAGCCCAAGGCTTCGAGGTC 5700  
5641 CACAGGCATGAAGGATATCATACCGCTGTGAGGCCAGAGCCCAAGGCTTCGAGGTC 5700  
5701 CACCTTCAACCTTCTACACCGCTTCCCGAGGTCGACATTTCCACCTGCACTGAGCC 5760  
5701 CACCTTCAACCTTCTACACCGCTTCCCGAGGTCGACATTTCCACCTGCACTGAGCC 5760  
5701 CACCTTCAACCTTCTACACCGCTTCCCGAGGTCGACATTTCCACCTGCACTGAGCC 5760  
5761 ACTGGGCGGACCCCTCATGAGGGGTCTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 5820  
5761 ACTGGGCGGACCCCTCATGAGGGGTCTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 5820  
5821 GAGAGCCCGCGGGGTGCGCCGCGCAAGCGGCGCCGAGCAACCGGCTCATGCTTCTCT 5880  
5821 GAGAGCCCGCGGGGTGCGCCGCGCAAGCGGCGCCGAGCAACCGGCTCATGCTTCTCTCT 5880  
5881 GCGCAAGCCCGAGCGGCTCGGGCTGAGAGCCGCTCTCTCCCAAGAGAGAGAGAGAGAGAGAG 5940  
5881 GCGCAAGCCCGAGCGGCTCGGGCTGAGAGCCGCTCTCTCCCAAGAGAGAGAGAGAGAGAGAG 5940  
5941 GCGCGGCGCCCTTACGCTCTGCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6000  
5941 GCGCGGCGCCCTTACGCTCTGCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6000  
6001 GAACCTGACCTTCAACAGCGCAAGCGGAGCCGCGCGGCACTGCTGCGCTCGGA 6060  
6001 GAACCTGACCTTCAACAGCGCAAGCGGAGCCGCGCGGCACTGCTGCGCTCGGA 6060  
6001 GAACCTGACCTTCAACAGCGCAAGCGGAGCCGCGCGGCACTGCTGCGCTCGGA 6060  
6061 CCGGACCGGGAGAAAGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 6120  
6061 CCGGACCGGGAGAAAGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 6120  
6121 TCTGGGTTACCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6180  
6121 TCTGGGTTACCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6180  
6181 CTACCCAGTGTGACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6240  
6181 CTACCCAGTGTGACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6240  
6181 CTACCCAGTGTGACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6240  
6241 CCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6300  
6241 CCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6300  
6301 CCGCCACCTCCCACTGCGGCGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6360  
6301 CCGCCACCTCCCACTGCGGCGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6360  
6361 CCAAGCCGCGCGGAGGATCAAGGATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6420  
6361 CCAAGCCGCGCGGAGGATCAAGGATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6420  
6421 TGAAGTATACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6480  
6421 TGAAGTATACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6480  
6481 GCGCCCGCTTCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6540  
6481 GCGCCCGCTTCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6540  
6481 GCGCCCGCTTCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6540  
6541 CAGTACCTTCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6600  
6541 CAGTACCTTCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6600  
6541 CAGTACCTTCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6600  
6601 CCAAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6660





; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq Version 4.0
; SEQ ID NO 653
; LENGTH: 8667
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-653

Query Match 98.8%; Score 8459.2; DB 5; Length 8667;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 8512; Conservative 1; Mismatches 38; Indels 11; Gaps 2;

QY 1 CATGTGCGGCTTCACACACAGCTTTGCGCACAGAGTGGAGGGCCACTGAGCCCGCTTACC 60
DB 115 CATGTGCGGATCCACACAGCTGTGGCACAGAGTGGAGGGCCACTGAGCCCGCTTACC 174
QY 61 GCCCCACAGCCTTTCTTACCCAGTGCAGATCGCCCCGAGACGACAGAGAGTGGGCTCTT 120
DB 175 GCCCCACAGCCTTTCTTACCCAGTGCAGATCGCCCCGAGACGACAGAGAGTGGGCTCTT 234
QY 121 GGAAGTACAGACACACTCCCGGACATATGCTCCACCTGTGCGCGGGCTCATCATCCA 180
DB 235 GGAAGTACAGACACACTCCCGGACATATGCTCCACCTGTGCGCGGGCTCATCATCCA 294
QY 181 GCCCAGCGCGGAGGCGCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCA 240
DB 295 GCCCAGCGCGGAGGCGCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCA 354
QY 241 GGAAGTCCACCTCGCGCCAGAGTCCCACTCATATCTGCGCCAGCTGGGGAAATCAGAGAT 300
DB 355 GGAAGTCCACCTCGCGCCAGAGTCCCACTCATATCTGCGCCAGCTGGGGAAATCAGAGAT 414
QY 301 GGAAGTTCATTGAAAGCAAGCGGCTCGGCTAGAGCTGCTGACCCCGCTGCGGACC 360
DB 415 GGAAGTTCATTGAAAGCAAGCGGCTCGGCTAGAGCTGCTGACCCCGCTGCGGACC 474
QY 361 GTCAACCCCTGTGCGCACAGGGCCAGCTGCGGAGTCTGAAAGACTCACCAAGAACCGTAG 420
DB 475 GTCAACCCCTGTGCGCACAGGGCCAGCTGCGGAGTCTGAAAGACTCACCAAGAACCGTAG 534
QY 421 CCTGACCGGCGAAGCTGGAACCGGTGTCTTCCCCAGCCCCCGCACACTGACCTGAGCT 480
DB 535 CCTGACCGGCGAAGCTGGAACCGGTGTCTTCCCCAGCCCCCGCACACTGACCTGAGCT 594
QY 481 GGAAGCTGGTGCAGCGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGAACCGGTGGA 540
DB 595 GGAAGCTGGTGCAGCGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGAACCGGTGGA 654
QY 541 CCGAGAGATCAACATGTGTAGAGCAGACAGATCTTAACTGAAAGAAAGACAGCAACAGCT 600
DB 655 CCGAGAGATCAACATGTGTAGAGCAGACAGATCTTAACTGAAAGAAAGACAGCAACAGCT 714
QY 601 GAGAGAGAGAGGCTGCGCAAGCGCGCCGAGCTGTGAAAGACCGCGTGTCAACGCGCCATGGA 660
DB 715 GAGAGAGAGAGGCTGCGCAAGCGCGCCGAGCTGTGAAAGACCGCGTGTCAACGCGCCATGGA 774
QY 661 GTGGAAGACCGGCAAGCTGTGTGACATCATCAAGACGGAACCGGAAGAAAGGCTGAAGC 720
DB 775 GTGGAAGACCGGCAAGCTGTGTGACATCATCAAGACGGAACCGGAAGAAAGGCTGAAGC 834
QY 721 TGCAATCGGATTTGTGAAGGCTTGGGGCCCGAGGTGAGCTGCGCTGTACAACAAGCC 780
DB 835 TGCAATCGGATTTGTGAAGGCTTGGGGCCCGAGGTGAGCTGCGCTGTACAACAAGCC 894
QY 781 CTCGCAACACCGGCGAGTATCATGAGAACTCAAAATTAACCAAGGCGATGCGGAAGAGCT 840
DB 895 CTCGCAACACCGGCGAGTATCATGAGAACTCAAAATTAACCAAGGCGATGCGGAAGAGCT 954
QY 841 AATCTTTGACTTCAAGAGAGGAATCAAGCTTGGAAACAAATGGAAGCAGAAATTTTGCCA 900
DB 955 AATCTTTGACTTCAAGAGAGGAATCAAGCTTGGAAACAAATGGAAGCAGAAATTTTGCCA 1014

QY 901 GCGCTATGACCAAGCTCATGAGAGCCTTGGAAAAAAGGTGAGGCCATCGAAAACAACCC 960
DB 1015 GCGCTATGACCAAGCTCATGAGAGCCTTGGAAAAAAGGTGAGGCCATCGAAAACAACCC 1074
QY 961 GCGCGCGCGGCGCAAGAGAGAGAGTGTGCGAGTACTACGAAGAGCATTCCTCGAGAT 1020
DB 1075 CCGGCGCGGCGCAAGAGAGAGAGTGTGCGAGTACTACGAAGAGCATTCCTCGAGAT 1134
QY 1021 CCGCAAGACGCGGACGCTGACAGAGCGCATGCGACAGAGGCTGGGCGACGCGGGCAGTGG 1080
DB 1135 CCGCAAGACGCGGACGCTGACAGAGCGCATGCGACAGAGGCTGGGCGACGCGGGCAGTGG 1194
QY 1081 GCTGTCCATGTGTGCGCGCGCGCGGACGACGAGAGTGTCAAGATCATCATGTGGCTCTC 1140
DB 1195 GCTGTCCATGTGTGCGCGCGCGGACGACGAGAGTGTCAAGATCATCATGTGGCTCTC 1254
QY 1141 AGAGCAGAGAAACCTGGAAGACAGATGCGCAGCTGACCGGATCCGCGCATCTGTGA 1200
DB 1255 AGAGCAGAGAAACCTGGAAGACAGATGCGCAGCTGACCGGATCCGCGCATCTGTGA 1314
QY 1201 CGACGCTGACAGACGCGCATCAAGTTTCATCAACATGAACGGGCTTATGCGCGACCCCAT 1260
DB 1315 CGACGCTGACAGACGCGCATCAAGTTTCATCAACATGAACGGGCTTATGCGCGACCCCAT 1374
QY 1281 GAAGGTGTACAAAGACCGCCAGGTTCATGAACATGTGAGTGAACAGAGAAAGAGACTT 1320
DB 1375 GAAGGTGTACAAAGACCGCCAGGTTCATGAACATGTGAGTGAACAGAGAAAGAGACTT 1434
QY 1321 CCGGGAAGAAAGTTCAATGACAGATCCCAAGACTTTGGCCGATGCGATCTTGTGAGAG 1380
DB 1435 CCGGGAAGAAAGTTCAATGACAGATCCCAAGACTTTGGCCGATGCGATCTTGTGAGAG 1494
QY 1381 GAAGCAGTGTGCTGAGTGCCTCTTATTACTACTGTACTGAAGAAATGAGAACTATTA 1440
DB 1495 GAAGCAGTGTGCTGAGTGCCTCTTATTACTACTGTACTGAAGAAATGAGAACTATTA 1554
QY 1441 GAGCTGTGTAGACGGAAGCTATTCGGCGCGCGCAAGGCCAGCAGCAACAACAGACGA 1500
DB 1555 GAGCTGTGTAGACGGAAGCTATTCGGCGCGCGCAAGGCCAGCAGCAACAACAGACGA 1614
QY 1501 GCAAGAGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560
DB 1615 GCAAGAGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1674
QY 1561 AGATGAGAAAGAGAAAGAAAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB 1675 AGATGAGAAAGAGAAAGAAAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1734
QY 1621 CGACAAAGGAAGACTCTCTCAAGAGAGAGACAGACGACTTCAAGGGAGAGCAACGACGA 1680
DB 1735 CGACAAAGGAAGACTCTCTCAAGAGAGAGACAGACGACTTCAAGGGAGAGCAACGACGA 1794
QY 1681 GAAAGAGGCTGTGGCTTCCAAAGGCGCGCAAACTGCGCAACGCGGAGAAAGCGAAAGG 1740
DB 1795 GAAAGAGGCTGTGGCTTCCAAAGGCGCGCAAACTGCGCAACGCGGAGAAAGCGAAAGG 1854
QY 1741 CCGCATCAACCGGCTCAATGTGCTTAATGAGGCCAACAGGAGAGGCTATCAACCCCAACA 1800
DB 1855 CCGCATCAACCGGCTCAATGTGCTTAATGAGGCCAACAGGAGAGGCTATCAACCCCAACA 1914
QY 1801 GAGCGCCGAGGCTGGCTTCATGAGAGCTGAATGAGATTCTGCTGTGACAGAGAAAGAAAT 1860
DB 1915 GAGCGCCGAGGCTGGCTTCATGAGAGCTGAATGAGATTCTGCTGTGACAGAGAAAGAAAT 1974
QY 1861 GGAACACGCGCAAGAAAGTCTTCTGTGAACAAGCGCGCAACTGTGTGCGCATTCGCCGAT 1920
DB 1975 GGAACACGCGCAAGAAAGTCTTCTGTGAACAAGCGCGCAACTGTGTGCGCATTCGCCGAT 2034
QY 1921 GGTGGGCTTCAAGACTGTGTGCGCAAGTGAAGCTTCACTTCAACTCAAGAAAGGCA 1980
DB 2035 GGTGGGCTTCAAGACTGTGTGCGCAAGTGAAGCTTCACTTCAACTCAAGAAAGGCA 2094
QY 1981 GAACCTCGATGAGATCTTGCAGACAGCAACGCTGAAGATGGAAGAGAGAGACGCGCG 2040

Db 2095 GAACTCGATGATCTTGACAGCACAAGCTGAAGATGAGAAAGAGAACGCGCG 2154  
Qy 2041 GAGGAAGAAAGAAAGCCGCGCGCGCGCGCAAGCGAGAGGCTGCAATCCCGCCCTGGT 2100  
Db 2155 GAGGAAGAAAGAAAGCCGCGCGCGCGCGCAAGCGAGAGGCTGCAATCCCGCCCTGGT 2214  
Qy 2101 GAGAGATGAGAGATGAGAGGCGCTCGGCGCTGAGCGGAAATGAGAGAGATGGTGGAGA 2160  
Db 2215 GAGAGATGAGAGATGAGAGGCGCTCGGCGCTGAGCGGAAATGAGAGAGATGGTGGAGA 2274  
Qy 2161 GGCTGAAGCTTATCATGCTCTTGGAATGAGGTGCCAAGGGGAAATGCAGTGCGCACG 2220  
Db 2275 GGCTGAAGCTTATCATGCTCTTGGAATGAGGTGCCAAGGGGAAATGCAGTGCGCACG 2334  
Qy 2221 CATCTGCAACAACAGCTCAAGACCGGAGCAATCCCTCTCTCCCAACTGAGGCGCCCAA 2280  
Db 2335 CATCTGCAACAACAGCTCAAGACCGGAGCAATCCCTCTCTCCCAACTGAGGCGCCCAA 2394  
Qy 2281 GGAACAAGGGCAGAAATGGGCGCCCAAGCGCCAGCACTTGAGCGCGGACGAGCCCGC 2340  
Db 2395 GGAACAAGGGCAGAAATGGGCGCCCAAGCGCCAGCACTTGAGCGCGGACGAGCCCGC 2454  
Qy 2341 AGGCCCCACCCCAACAAGAGACATCCGAGGCCCCCATTTGAGCCCAACCCCGGCTTC 2400  
Db 2455 AGGCCCCACCCCAACAAGAGACATCCGAGGCCCCCATTTGAGCCCAACCCCGGCTTC 2514  
Qy 2401 TGAAGCAACCGGAGCGCCCTTACGCCCCCAACAAGACCCCACTGCTGCACTCTTCC 2460  
Db 2515 TGAAGCAACCGGAGCGCCCTTACGCCCCCAACAAGACCCCACTGCTGCACTCTTCC 2574  
Qy 2461 TGTGTGCCCCAAGAGAGAGAGAGAGAGACCGGAGCGCGCCCACTGAGAGAGAG 2520  
Db 2575 TGTGTGCCCCAAGAGAGAGAGAGAGAGAGACCGGAGCGCGCCCACTGAGAGAGAG 2634  
Qy 2521 GAGAGAGCAAGAGCGCCCGCGGCTTGAGAGCTGCGCATGGAACAAGGAAAGCCAGAGA 2580  
Db 2635 GAGAGAGCAAGAGCGCCCGCGGCTTGAGAGCTGCGCATGGAACAAGGAAAGCCAGAGA 2694  
Qy 2581 GCCCGTCAGAGAGAGATGACCGGAGAGAACCGAGAGGGGCGCGCAAGGCAAGACCG 2640  
Db 2695 GCCCGTCAGAGAGAGATGACCGGAGAGAACCGAGAGGGGCGCGCAAGGCAAGACCG 2754  
Qy 2641 GAGAGCGGCTGAGGCGCACGCGGAGGGGCGCTCAAGGCAAGAAAGAGAGGCGGAG 2700  
Db 2755 GAGAGCGGCTGAGGCGCACGCGGAGGGGCGCTCAAGGCAAGAAAGAGAGGCGGAG 2814  
Qy 2701 CGGCAAGGCGCACACTGCGCAAGAGCTCGGCGCGCCCCCAGAGACGCGACTCCAGTGTAC 2760  
Db 2815 CGGCAAGGCGCACACTGCGCAAGAGCTCGGCGCGCCCCCAGAGACGCGACTCCAGTGTAC 2874  
Qy 2761 CTGCAAGTGCAGACGAGGTGATGAGGCGGAGGCGCGGCAAGAACCGGCTGTGTCCC 2820  
Db 2875 CTGCAAGTGCAGACGAGGTGATGAGGCGGAGGCGCGGCAAGAACCGGCTGTGTCCC 2934  
Qy 2821 AAGGCGCAGGCGCTCAACCGGACTGGGAGACCCCGGCGCAATGCGCTCAACCGAGAGCG 2880  
Db 2935 AAGGCGCAGGCGCTCAACCGGACTGGGAGACCCCGGCGCAATGCGCTCAACCGAGAGCG 2994  
Qy 2881 ACTGCACTGAAGAGCTGGAACGAGCGAGCGGCTGCGCATCCCGCCCACTCAGGTCAACAA 2940  
Db 2995 ACTGCACTGAAGAGCTGGAACGAGCGAGCGGCTGCGCATCCCGCCCACTCAGGTCAACAA 3054  
Qy 2941 AGTCCATGAGCGCGCGCGGAGAGACGAGCTCCCAACAGCGAGCTCCCGCAGCGCCACG 3000  
Db 3055 AGTCCATGAGCGCGCGCGGAGAGACGAGCTCCCAACAGCGAGCTCCCGCAGCGCCACG 3114  
Qy 3001 GGCACCGCAAAACCTGCAAGCGGAGAGAGCGCCCTCTGACAGCGCTGGAGAGCGCGCG 3060  
Db 3115 GGCACCGCAAAACCTGCAAGCGGAGAGAGCGCCCTCTGACAGCGCTGGAGAGCGCGCG 3174  
Qy 3061 GGGCAAGAGCAAGAGCGCGCGCAACCCCGCGCGCAAGAGAGGCTTGCAGCGCGAGCGCA 3120

Db 3175 GGGCAAGAGCAAGAGCGCGCGCAACCCCGCGCGCAAGAGAGGCTTGCAGCGCGAGCGCA 3234  
Qy 3121 GAAGCTGCGGGGAGACCCGCTTGTGAGACTTCGCGGCTGGCTTCCGAGTGCAGCGCG 3180  
Db 3235 GAAGCTGCGGGGAGACCCGCTTGTGAGACTTCGCGGCTGGCTTCCGAGTGCAGCGCG 3294  
Qy 3181 TGAAGTATCAAGGCTTCCCGCATGCGCCGAGACCCCTCAGAGCTTCTCTACGCTCAC 3240  
Db 3295 TGAAGTATCAAGGCTTCCCGCATGCGCCGAGACCCCTCAGAGCTTCTCTACGCTCAC 3354  
Qy 3241 TGGTCAACCACTGCGCTTGGGCTCCATGACACTGCGCGGCGCTGCTGCGCGCGCAC 3300  
Db 3355 TGGTCAACCACTGCGCTTGGGCTCCATGACACTGCGCGGCGCTGCTGCGCGCGCAC 3414  
Qy 3301 CACCATCTCAACCGCGCTCCGCTCATCTGTCAGAGCAACCCAGAGGCTCTCGAGAG 3360  
Db 3415 CACCATCTCAACCGCGCTCCGCTCATCTGTCAGAGCAACCCAGAGGCTCTCGAGAG 3474  
Qy 3361 GCAAAATAGGTGCATCTCCCAAGAAATGTGCGTCCAGTTCAGTCCGTAATCAGAGCA 3420  
Db 3475 GCAAAATAGGTGCATCTCCCAAGAAATGTGCGTCCAGTTCAGTCCGTAATCAGAGCA 3534  
Qy 3421 TGCCAAGGCGCGGTGGGCGCTGTGTCATGAGGAGCTGCGCTGCGCTGAGATCCCAAAA 3480  
Db 3535 TGCCAAGGCGCGGTGGGCGCTGTGTCATGAGGAGCTGCGCTGCGCTGAGATCCCAAAA 3594  
Qy 3481 GCTGCAACCTTCAAGGAGTGAAGAGAGAGAGAGAGTCCCAAGGCGCGGCTGGGCG 3540  
Db 3595 GCTGCAACCTTCAAGGAGTGAAGAGAGAGAGAGTCCCAAGGCGCGGCTGGGCG 3654  
Qy 3541 ACCGAGAGGCTGGGAGGTGCCCAAGCGCCAGAGGCGTCCGTGCTGAGAGGACAGCTCT 3600  
Db 3655 ACCGAGAGGCTGGGAGGTGCCCAAGCGCCAGAGGCGTCCGTGCTGAGAGGACAGCTCT 3714  
Qy 3601 GGGCTCAGTTCCGGGCGAAGAGATCAACAAAGCAATTCCACAGCAACGGGTGCTTCGGA 3660  
Db 3715 GGGCTCAGTTCCGGGCGAAGAGATCAACAAAGCAATTCCACAGCAACGGGTGCTTCGGA 3774  
Qy 3661 CAGCGCATCAATACCGGCGTCCATCAACCGGCAAGCGGCAAGCTGCTGTGACAA 3720  
Db 3775 CAGCGCATCAATACCGGCGTCCATCAACCGGCAAGCGGCAAGCTGCTGTGACAA 3834  
Qy 3721 GGGCAACATCAACAGGATATCGGCGAGAGACAGCCGAGTCCGTTGAGACGCGCGGGA 3780  
Db 3835 GGGCAACATCAACAGGATATCGGCGAGAGACAGCCGAGTCCGTTGAGACGCGCGGGA 3894  
Qy 3781 GGAACAGCTTCCCAAGGCGCATGTACTTACAAAGGCAAGAGGCGCATGCTTGTCTTA 3840  
Db 3895 GGAACAGCTTCCCAAGGCGCATGTACTTACAAAGGCGCATGCTTGTCTTA 3954  
Qy 3841 TGAAGGTGCAATGTCTTGAACCAATGCTTCCAAAGAGAGCGGCAAGAGAGCTCAGGAGC 3900  
Db 3955 TGAAGGTGCAATGTCTTGAACCAATGCTTCCAAAGAGAGCGGCAAGAGAGCTCAGGAGC 4014  
Qy 3901 CCCCATGAGAGCGGCGCGCCCAAGCGGCACTATGACATGAGAGGCGCGGTGGGAG 3960  
Db 4015 CCCCATGAGAGCGGCGCGCCCAAGCGGCACTATGAGAGAGGCGCGGTGGGAG 4074  
Qy 3961 AGCCATTTCTTGAAGCAGAGATGAAGGTCTATGAGGCGGTGCATCCCGCGGAGGACA 4020  
Db 4075 AGCCATTTCTTGAAGCAGAGATGAAGGTCTATGAGGCGGTGCATCCCGCGGAGGACA 4134  
Qy 4021 CAGCCCCCAACCTTCAAGAGACAGACACATTCGCGGGGTCCATTCACAAAGGATCCC 4080  
Db 4135 CAGCCCCCAACCTTCAAGAGACAGACACATTCGCGGGGTCCATTCACAAAGGATCCC 4194  
Qy 4081 TCGGTCTTACGTGAGAGCAACAGAGGACTTACGTGAGTGGGAGGCGCAAGCTCTTAAAGG 4140  
Db 4195 TCGGTCTTACGTGAGAGCAACAGAGGACTTACGTGAGTGGGAGGCGCAAGCTCTTAAAGG 4254  
Qy 4141 GAGAGGCAACGCTTCCGCGCCCAACGCGCTTCAAGGAGCTTGAACGAGGCTTAAAGGCA 4200  
Db 4255 GAGAGGCAACGCTTCCGCGCCCAACGCGCTTCAAGGAGCTTGAACGAGGCTTAAAGGCA 4314

OY 4201 GGCCTTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAAGGCGCTGTGGCCACGGTGAAGA 4260  
Db 4315 GGCCTTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAAGGCGCTGTGGCCACGGTGAAGA 4374  
OY 4281 GGCCTTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAAGGCGCTGTGGCCACGGTGAAGA 4320  
Db 4375 GGCCTTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAAGGCGCTGTGGCCACGGTGAAGA 4434  
OY 4321 CTGGCCCCCGGGCGGCTCAAGAGGAGCTCATCAAGAGGCAACCCGCTCAAGTACGA 4380  
Db 4435 CTTGGCCCCCGGGCGGCTCAAGAGGAGCTCATCAAGAGGCAACCCGCTCAAGTACGA 4494  
OY 4381 CACCGGCGGCTCAACCATCTGCTCCAAAAAGCAGCGTACGCTCCCTCATCGGAGGCC 4440  
Db 4495 CACCGGCGGCTCAACCATCTGCTCCAAAAAGCAGCGTACGCTCCCTCATCGGAGGCC 4554  
OY 4441 CGGCGGGAAGTTCCCAACCCGCTGCAACCCGCTGGAATGATGGCCAGCCCGGCACTGGA 4500  
Db 4555 CGGCGGGAAGTTCCCAACCCGCTGCAACCCGCTGGAATGATGGCCAGCCCGGCACTGGA 4614  
OY 4501 ACGTCTCTCTACAGAGAGGCTGAAGAGCCGGCCAGGAGCCGCGCAGCAGCTCGGGGG 4560  
Db 4615 ACGTCTCTCTACAGAGAGGCTGAAGAGCCGGCCAGGAGCCGCGCAGCAGCTCGGGGG 4674  
OY 4561 CTCCATTGCGCGCGCGCGCCCGGTCAATTGTGCTGAAGCTGGTAAAGCCGCGCAGAGCCC 4620  
Db 4675 CTCCATTGCGCGCGCGCGCCCGGTCAATTGTGCTGAAGCTGGTAAAGCCGCGCAGAGCCC 4734  
OY 4621 CCGTGCCTATAGAGACACAGGGGCAACCTTTGCGCGCCACTTCCCAAGAGTTCCCGCT 4680  
Db 4735 CTTAACCCTATAGAGACACAGGGGCAACCTTTGCGCGCCACTTCCCAAGAGTTCCCGCT 4794  
OY 4681 GACCATGCGGAGGCCACGCGCGCGCTGCAAGAGGAGCCCTTTGCTCCAGAAAGCATC 4740  
Db 4795 GACCATGCGGAGGCCACGCGCGCGCTGCAAGAGGAGCCCTTTGCTCCAGAAAGCATC 4854  
OY 4741 CCAAGACCGAAAGCTGACGCTCGACGCTTGTGAGATTCGCCAAGTCCCGCAGACGCGT 4800  
Db 4855 CCAAGACCGAAAGCTGACGCTCGACGCTTGTGAGATTCGCCAAGTCCCGCAGACGCGT 4914  
OY 4801 GCCCGAGACCCACCCACACCCCATCTCGCGCTATAGAGCATCTGCTTCGGGGGCTGAGTGG 4860  
Db 4915 GCCCGAGACCCACCCACACCCCATCTCGCGCTATAGAGCATCTGCTTCGGGGGCTGAGTGG 4974  
OY 4861 CGTGAACCTGTATCGCAGGCAATGCCCCCTGAGCTTTCGACCCCACTTCATACCCCGAG 4920  
Db 4975 CGTGAACCTGTATCGCAGGCAATGCCCCCTGAGCTTTCGACCCCACTTCATACCCCGAG 5034  
OY 4921 CATCCCTTGAACGAGCGCGCTGCTACTACCTGCCCCGAGCATGAGCCCCCAACCCAC 4980  
Db 5035 CATCCCTTGAACGAGCGCGCTGCTACTACCTGCCCCGAGCATGAGCCCCCAACCCAC 5094  
OY 4981 CTACCCCGACCTGTATCCACCCCTACTACTACCGCGGCTACCCCGACAGGGCGGCTGGA 5040  
Db 5095 CTACCCCGACCTGTATCCACCCCTACTACTACCGCGGCTACCCCGACAGGGCGGCTGGA 5154  
OY 5041 GAAACGGAGACCATCATCATATGATCTACATCACTCGCAGAGATGACCAACAAGCGGC 5100  
Db 5155 GAAACGGAGACCATCATCATATGATCTACATCACTCGCAGAGATGACCAACAAGCGGC 5214  
OY 5101 CACCGCATGCGCCAGCGAGCTGATATGCTGAGGGGCGCTTGCACCAGAGTCTTCGCT 5160  
Db 5215 CACCGCATGCGCCAGCGAGCTGATATGCTGAGGGGCGCTTGCACCAGAGTCTTCGCT 5274  
OY 5161 GGCATCTCAACTACGCTGCGGGGTCCCGAGGATCATTCGACTGTCCCAAGTCCACACTT 5220  
Db 5275 GGCATCTCAACTACGCTGCGGGGTCCCGAGGATCATTCGACTGTCCCAAGTCCACACTT 5334  
OY 5221 GCGCTGTGCTGTCGCGCCGAGCAGCAGGACCCCAAGCAGCCGCTTGGCTGA 5280  
Db 5335 GCGCTGTGCTGTCGCGCCGAGCAGCAGGACCCCAAGCAGCCGCTTGGCTGA 5394

OY 5281 CTTCCCAACCGCGGCCCAACCCCTTCAAGACCGCGCCACAGCAGCTTCCCATCTTCCCAAG 5340  
Db 5395 CTTCCCAACCGCGGCCCAACCCCTTCAAGACCGCGCCACAGCAGCTTCCCATCTTCCCAAG 5454  
OY 5341 AGTTCCAACACACTTGAAGAAACCAACCAACAGCTTCGTGTCCAGGGGGAGCGAGACCG 5400  
Db 5455 AGTTCCAACACACTTGAAGAAACCAACCAACAGCTTCGTGTCCAGGGGGAGCGAGACCG 5514  
OY 5401 GGATGAGAGCGGAGCCGGAGATCGGGAGCGGGAAAGTCCATCTTCATCGTCCACACGAC 5460  
Db 5515 GGATGAGAGCGGAGCCGGAGATCGGGAGCGGGAAAGTCCATCTTCATCGTCCACACGAC 5574  
OY 5461 GGTGAGACACCAACCATCTTGAAGACTGTGTACAGAGACAGCAGCGCAGCAGCGGAG 5520  
Db 5575 GGTGAGACACCAACCATCTTGAAGACTGTGTACAGAGACAGCAGCGCAGCAGCGGAG 5634  
OY 5521 CAGCGCGGGGTGGGGGACAGAGCGCGCGCGCTCCCATCTCCATATGCCACACAGCA 5580  
Db 5635 CAGCGCGGGGTGGGGGACAGAGCGCGCGCGCTCCCATCTCCATATGCCACACAGCA 5694  
OY 5581 CTGCGCCATCTCCCTCGGACCCAGAGATGCCCTCCAGCAGAGACCCAGTGTTCACAA 5640  
Db 5695 CTGCGCCATCTCCCTCGGACCCAGAGATGCCCTCCAGCAGAGACCCAGTGTTCACAA 5754  
OY 5641 CACAGCATGAAGGATCATACCGCTGTGAGAGCCACAGCAAGCCCATGCTTGAAGTTC 5700  
Db 5755 CACAGCATGAAGGATCATACCGCTGTGAGAGCCACAGCAAGCCCATGCTTGAAGTTC 5814  
OY 5701 CACCTCCACCTTCACACCGGTTGCGCGCAGCTGCCACATTCACACCTGCAACCACTGCC 5760  
Db 5815 CACCTCCACCTTCACACCGGTTGCGCGCAGCTGCCACATTCACACCTGCAACCACTGCC 5874  
OY 5761 ACTGGCGGCAACCTTCGATGGGGTCTACCTTACCTCATGAGCCCGTCTTGTGCGCAA 5820  
Db 5875 ACTGGCGGCAACCTTCGATGGGGTCTACCTTACCTCATGAGCCCGTCTTGTGCGCAA 5934  
OY 5821 GAGAGCCCCCGGGGTGCGCGCGCAGAGCGGCCCGAGACACACCGGCAATGCTTCTCT 5880  
Db 5935 GAGAGCCCCCGGGGTGCGCGCGCAGAGCGGCCCGAGACACACCGGCAATGCTTCTCT 5994  
OY 5881 GCGCAAGCCCCAGCGGCTCGGGCTGGAAGCCGCGCTCTCCCGACAGAGGCTCGGA 5940  
Db 5995 GCGCAAGCCCCAGCGGCTCGGGCTGGAAGCCGCGCTCTCCCGACAGAGGCTCGGA 6054  
OY 5941 GCCCGGCCCCCTAGTGCCTCTGTCTCTGAGCAGGCAACATTCGCCGACACCCCTGCGAA 6000  
Db 6055 GCCCGGCCCCCTAGTGCCTCTGTCTCTGAGCAGGCAACATTCGCCGACACCCCTGCGAA 6114  
OY 6001 GAACCTGACCTTCAACAGCGCAAGCCGAGCCGCGCGCGCACTGCTTGGCTCGGA 6060  
Db 6115 GAACCTGACCTTCAACAGCGCAAGCCGAGCCGCGCGCGCACTGCTTGGCTCGGA 6174  
OY 6061 CCGCAGCGGGGAAAGACTCAAGATTAACCTTTTCCATCCAGGAACGTGGAACCTCCGTTTC 6120  
Db 6175 CCGCAGCGGGGAAAGACTCAAGATTAACCTTTTCCATCCAGGAACGTGGAACCTCCGTTTC 6234  
OY 6121 TCTGGTTTACACGAGCAGCACTACAGCCCGGAAGGGGTGAGGCCCTGCTGAG 6180  
Db 6235 TCTGGTTTACACGAGCAGCACTACAGCCCGGAAGGGGTGAGGCCCTGCTGAG 6294  
OY 6181 CTACCCAGTGTGACCCACAGCAAGGGGCTCCCAAGACCTTGAAGAGCTTGCACAAAG 6240  
Db 6295 CTACCCAGTGTGACCCACAGCAAGGGGCTCCCAAGACCTTGAAGAGCTTGCACAAAG 6354  
OY 6241 CCACCTGAGGGGGAGCTGCGGCGCCAGCAAGCCCGCTGAAAGCTTGGCGGGGAGGC 6300  
Db 6355 CCACCTGAGGGGGAGCTGCGGCGCCAGCAAGCCCGCTGAAAGCTTGGCGGGGAGGC 6414  
OY 6301 GCGCCACTTCCCAACCTTGGCGGCTGCTGTGAGAGCAGGCTTGTTCAGCGCGCTGCT 6360  
Db 6415 GCGCCACTTCCCAACCTTGGCGGCTGCTGTGAGAGCAGGCTTGTTCAGCGCGCTGCT 6474  
OY 6361 CCAAGCGGCCCCAGGGGTCAAAAGTCAACAGCGGGGTGTCACCTTGGCCAGCAGCATAG 6420

6475 CCAAGACCGCCCGGAGGCTCAAAAGTCAACGACGGGTGTCAACCTGAGCCACGACATCAG 6534  
6421 TGAAGGTATCAACAACAGACTACACCGGGACCAACCAAGCAAGCTAGAGGACCCCTGAC 6480  
6535 TGAAGGTATCAACAACAGACTACACCGGGACCAACCAAGCAAGCTAGAGGACCCCTGAC 6594  
6481 CCGCCCCCTCTACTCTCTCCCTGAGGACAGCTGCCCCGTCTTGGAACCTCCGCGGCCAC 6540  
6595 CCGCCCCCTCTACTCTCTCCCTGAGGACAGCTGCCCCGTCTTGGAACCTCCGCGGCCAC 6654  
6541 CAGTGACTCTTACTCTCCGCGCGGACCAATGATGCCCGGCGCTGAGCTCCCCCAAG 6600  
6655 CAGTGACTCTTACTCTCCGCGCGGACCAATGATGCCCGGCGCTGAGCTCCCCCAAG 6714  
6601 CCAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGCTGGTTGGTGGTGGTGAAGA 6660  
6715 CCAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGCTGGTTGGTGGTGGTGAAGA 6774  
6661 CCGTATTTGAACCTGTGCTCCCAACCGAGGGCATGACGAGCCAGGCACTCCCGAGTGC 6720  
6775 CCGTATTTGAACCTGTGCTCCCAACCGAGGGCATGACGAGCCAGGCACTCCCGAGTGC 6834  
6721 TGTGTACCCGCTGCTGTACCGGAGTGGGAAAGACGAGGCCAGAGATGGGCTTCAA 6780  
6835 TGTGTACCCGCTGCTGTACCGGAGTGGGAAAGACGAGGCCAGAGATGGGCTTCAA 6894  
6781 GTCTCCAGGCAACACAGGACGCGGACGCTTCTTACAGAGCTGACCGAGGCAATC 6840  
6895 GTCTCCAGGCAACACAGGACGCGGACGCTTCTTACAGAGCTGACCGAGGCAATC 6954  
6841 CCGCATGTGCAAGTCCAGAGGCAAGATCAACAAGAGCTGAACCCACAACCGGAA 6900  
6955 CCGCATGTGCAAGTCCAGAGGCAAGATCAACAAGAGCTGAACCCACAACCGGAA 7014  
6901 TGAAGCTGAATACAATATAGCCAGGCTGGAGCGAGATCTTCAATATGCCCGCATAC 6960  
7015 TGAAGCTGAATACAATATAGCCAGGCTGGAGCGAGATCTTCAATATGCCCGCATAC 7074  
6961 CCGAACAAGGCTTTATGACCTTATGAAGCCAGAGCGGTGAGGAACATGCCAGCAACAT 7020  
7075 CCGAACAAGGCTTTATGACCTTATGAAGCCAGAGCGGTGAGGAACATGCCAGCAACAT 7134  
7021 GGGGCTGAGGCGCATATATGAAGGCACTCATGGTAAATATGACAGTGGGAGAAGTC 7080  
7135 GGGGCTGAGGCGCATATATGAAGGCACTCATGGTAAATATGACAGTGGGAGAAGTC 7194  
7081 CCGCGCGCTCAGGCGCCATGCTTTTAAACCTCTGAATGCCAGTGCAGGCTGCCGCTGC 7140  
7195 CCGCGCGCTCAGGCGCCATGCTTTTAAACCTCTGAATGCCAGTGCAGGCTGCCGCTGC 7254  
7141 TATGCCCATTAACCGCTGCTGACGGAAGATGACCACTCACTCCGCAAGTGGCGG 7200  
7255 TATGCCCATTAACCGCTGCTGACGGAAGATGACCACTCACTCCGCAAGTGGCGG 7314  
7201 CCGGAAGGCGCAAGTCTCTGAGGACCGGACGCGGAAAGCCAAATCCCGGCGCGG 7260  
7315 CCGGAAGGCGCAAGTCTCTGAGGACCGGACGCGGAAAGCCAAATCCCGGCGCGG 7374  
7261 CCTGGCATCTGGGAGACCGGCAACCTCTGTCTCTCACTGACACTCGAGGAGACTGCAA 7320  
7375 CCTGGCATCTGGGAGACCGGCAACCTCTGTCTCTCACTGACACTCGAGGAGACTGCAA 7434  
7321 CCGCGGAGAGCGGCTCAACCAACGCGGTGAGGAGACAGGCTTGTCCGAGGTTCCAC 7380  
7435 CCGCGGAGAGCGGCTCAACCAACGCGGTGAGGAGACAGGCTTGTCCGAGGTTCCAC 7494  
7381 GGCATTCCCTCAACCCCTCTGATCATGCGGCTGACGAGGCGGTGTATGGCTTCCAC 7440  
7495 GGCATTCCCTCAACCCCTCTGATCATGCGGCTGACGAGGCGGTGTATGGCTTCCAC 7554  
7441 CCGCACGAGGCTTCCCGCGGACGCGGCGCTTGTGCGCCCAACAGGCTGGAGCA 7500

7555 CCCACGAGGCTTCCCGCGGAGAGCGGCGCTTGTGCGCCCAACAGGCTGGAGCA 7614  
7501 GAGGCCAAGCACTGCTGTGCTGAGTACGAGCACTCTCCAGACAGGAGTACTCAG 7560  
7615 GAGGCCAAGCACTGCTGTGCTGAGTACGAGCACTCTCCAGAGGAGGAGTACTCAG 7674  
7561 AACAGGCGGGGGGGGGCGGGCGGTBUBERS PERLUGTCAAGTCCAGGACCAAGAA 7620  
7675 AACAGGCGGGGGGGGGGGCGG-----TGTCAAGTCCAGGACCAAGAA 7724  
7621 CCGGCTTGAAGAGCGGGCGGCTGCGCACTCCCCCAACAAGGAAGAGCCCTAGTC 7680  
7725 CCGGCTTGAAGAGCGGGCGGCTGCGCACTCCCCCAACAAGGAAGAGCCCTAGTC 7784  
7681 CCGCTGCGCTCAATCATCTGTGCTGAGAGCCGCAATCTTGTGCTGTAAAGCTT 7740  
7785 CCGCTGCGCTCAATCATCTGTGCTGAGAGCCGCAATCTTGTGCTGTAAAGCTT 7844  
7741 AACTAAGACTCCCGCCCGGAGCTGGCCCTGTGCAAGCTTACTCAGGGAGTGTTA 7800  
7845 AACTAAGACTCCCGCCCGGAGCTGGCCCTGTGCAAGCTTACTCAGGGAGTGTTA 7904  
7801 GTGCTGGGAAGGAGGGGAAGGGGCTCGGGAGGGGGCAACGAGGCTGTGCAAGCA 7860  
7905 GTGCTGGGAAGGAGGGGAAGGGGCTCGGGAGGGGGCAACGAGGCTGTGCAAGCA 7964  
7861 ACAAGGGGGCGAGGGCGGCGAGGAGCCAAAGCAGATGACCAAGCACTCCACAGCCAC 7920  
7965 ACAGAGGGCGCAAGGGCGGCGAGGAGCCAAAGCAGATGACCAAGCACTCCACAGCCAC 8024  
7921 TGCCTCCCGGAATGATTTGGAACCAAGTCTAACTGAGCTCGAGGCCCGCGCT 7980  
8025 TGCCTCCCGGAATGATTTGGAACCAAGTCTAACTGAGCTCGAGGCCCGCGCT 8084  
7981 CCTCTCGCTCCCATCTCCGCTTAAGGCTGTGACAGATGAGCGAGGCCCTGTGCAAGCC 8040  
8085 CCTCTCGCTCCCATCTCCGCTTAAGGCTGTGACAGATGAGCGAGGCCCTGTGCAAGCC 8144  
8041 CCAAGTGGCTGTTCCGGTCCCGACCAAGCTGCCCAAGCAAGATTTCTGGAACCA 8100  
8145 CCAAGTGGCTGTTCCGGTCCCGACCAAGCTGCCCAAGCAAGATTTCTGGAACCA 8204  
8101 GTCAAGGCGAGTGGGCGGACAAAGGGCGAGGTGCGGCTGGGGGGGAACGATGCTCCGA 8160  
8205 GTCAAGGCGAGTGGGCGGACAAAGGGCGAGGTGCGGCTGGGGGGGAACGATGCTCCGA 8264  
8161 GGAAGTGAAGTGTGTTTTCACATCTGTTGCGGACGCGGTGGGAAGAAAGCAATGTA 8220  
8265 GGAAGTGAAGTGTGTTTTCACATCTGTTGCGGACGCGGTGGGAAGAAAGCAATGTA 8324  
8221 AATGATGCTGTGGTTTACAGGGATATTTTATGATACCTTCAATGAATTAATCAGATGTT 8280  
8325 AATGATGCTGTGGTTTACAGGGATATTTTATGATACCTTCAATGAATTAATCAGATGTT 8384  
8281 TTAAGCAAGAAAGACTTACCAAGTATTAAGTCTGTGCTTTTATGATCTGTCTTAACG 8340  
8385 TTAAGCAAGAAAGACTTACCAAGTATTAAGTCTGTGCTTTTATGATCTGTCTTAACG 8444  
8341 TTAAGCAAGGCTGTGAGGCGGACAGTGGTGAACCCATCACTCGAGGACCAAGGGGCG 8400  
8445 TTAAGCAAGGCTGTGAGGCGGACAGTGGTGAACCCATCACTCGAGGACCAAGGGGCG 8504  
8401 GGGGACTGCTCG-TCAAGCGCGCGGTGTGCTCCCTCCCTCTTGGGAGAAATGA 8459  
8505 GGGGACTGCTCG-TCAAGCGCGCGGTGTGCTCCCTCCCTCTTGGGAGAAATGA 8564  
8460 ATTTCAGTGGTATTTCTGTGCGGCAATTTGCGAGGAGTGGTATTTCTGTCAATTAAC 8519  
8565 ATTTCAGTGGTATTTCTGTGCGGCAATTTGCGAGGAGTGGTATTTCTGTCAATTAAC 8624  
8520 ACGTGTCTTAATTAATAAGCAATTAATCTCCAAAAA 8561  
8625 ACGTGTCTTAATTAATAAGCAATTAATCTCCAAAA 8666

## RESULT 4

US-09-819-104A-1

; Sequence 1, Application US/09819104A

; Publication No. US20030027137A1

; GENERAL INFORMATION:

; APPLICANT: Chem. J. Don

; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES

; FILE REFERENCE: UMG-030

; CURRENT APPLICATION NUMBER: US/09/819,104A

; PRIOR FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 8686

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (157)..(7677)

; US-09-819-104A-1

ABN

Query Match 97.2%; Score 8324.6; DB 3; Length 8686;

Best Local Similarity 98.6%; Pred. No. 0;

Matches 8472; Conservative 1; Mismatches 34; Indels 88; Gaps 4;

QY 1 CATGTGGGCTTCACACAGCTTTGTG3CACAGCTGAGAGGCGCACTGAGCCCGCTACCC 60  
DB 156 CATGTGGGCTTCACACAGCTGTG3CACAGCTGAGAGGCGCACTGAGCCCGCTACCC 215  
QY 61 GCGCCACAGCCCTTCTCTCCAGTGACAGTGCCTCGGACGCAACAGGACGTGCGGCTCT 120  
DB 216 GCGCCACAGCCCTTCTCTCCAGTGACAGTGCCTCGGACGCAACAGGACGTGCGGCTCT 275  
QY 121 GAGATGACAGACCACTCCCGGCACTATGCTCCGACCTGTGCGGCGCTCATCATCA 180  
DB 276 GAGATGACAGACCACTCCCGGCACTATGCTCCGACCTGTGCGGCGCTCATCATCA 335  
QY 181 GCGCCAGCGGCGGAGGCGCTCTCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCA 240  
DB 336 GCGCCAGCGGCGGAGGCGCTCTCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCA 395  
QY 241 GAGCTTCAACCTGCGGCGGAGTCCCACTCATCTGCGGCGGAGTGAAGAT 300  
DB 396 GAGCTTCAACCTGCGGCGGAGTCCCACTCATCTGCGGCGGAGTGAAGAT 455  
QY 301 GAGATTATTGAAACAAAGCGCCTCGGCTAGAGCTGCTGACCCCTGCTGACG 360  
DB 456 GAGATTATTGAAACAAAGCGCCTCGGCTAGAGCTGCTGACCCCTGCTGACG 515  
QY 361 GTACCCCTGCTGCGGCGGAGTCCCACTCATCTGCGGCGGAGTGAAGAT 420  
DB 516 GTACCCCTGCTGCGGCGGAGTCCCACTCATCTGCGGCGGAGTGAAGAT 575  
QY 421 CCGTACCGGCGGAGTGAAGCGGTTCTCCCGGAGCGGCGGCGGAGTGAAGAT 480  
DB 576 CCGTACCGGCGGAGTGAAGCGGTTCTCCCGGAGCGGCGGCGGAGTGAAGAT 635  
QY 481 GAGCTGTGCGGCGGAGTGAAGCGGTTCTCCCGGAGCGGCGGCGGAGTGAAGAT 540  
DB 636 GAGCTGTGCGGCGGAGTGAAGCGGTTCTCCCGGAGCGGCGGCGGAGTGAAGAT 695  
QY 541 CCGAGAGATCAACATGTGAGAGCAGATCTTAAGCTGAAGAGAGCAGCAAGCT 600  
DB 696 CCGAGAGATCAACATGTGAGAGCAGATCTTAAGCTGAAGAGAGCAGCAAGCT 755  
QY 601 GAGAGAGAGAGCTGCGGCGGAGCGGCGGCGGAGTGAAGATCGGCGGCGGAGTGAAG 660  
DB 756 GAGAGAGAGAGCTGCGGCGGAGCGGCGGCGGAGTGAAGATCGGCGGCGGAGTGAAG 815

QY 661 GTGAGAGACCGCAGCGCTGTGAGATCATCAGACGAGAACCGGAGAGAGCTGAAGC 720  
DB 816 GTGAGAGACCGCAGCGCTGTGAGATCATCAGACGAGAACCGGAGAGAGCTGAAGC 875  
QY 721 TGCACATGAGATTGTGAGAGCGCTGAGGCGCCAGGTGAGCTGCGGCTGACACCAAGC 780  
DB 876 TGCACATGAGATTGTGAGAGCGCTGAGGCGCCAGGTGAGCTGCGGCTGACACCAAGC 935  
QY 781 CTCGACACCGGCGGAGTATGAGAACATCAAAATAACGAGGATGCGGAGAGACT 840  
DB 936 CTCGACACCGGCGGAGTATGAGAACATCAAAATAACGAGGATGCGGAGAGACT 995  
QY 841 AATCTTGAATCAAGAGAGAAATCAAGCTCGGAGAAATGAGAGCAAGATTCTGCA 900  
DB 996 AATCTTGAATCAAGAGAGAAATCAAGCTCGGAGAAATGAGAGCAAGATTCTGCA 1055  
QY 901 GCGCTATGACCAAGCTCATGAGAGCGCTTGAAGAAAAAGGTGAGCGCATCGAAGAACCC 960  
DB 1056 GCGCTATGACCAAGCTCATGAGAGCGCTTGAAGAAAAAGGTGAGCGCATCGAAGAACCC 1115  
QY 961 GCGCGCGGCGGCGAAGAGAGCAAGTGCAGAGTACTAGAGAAAGCAAGTCCCTGAGAT 1020  
DB 1116 CCGGCGGCGGCGAAGAGAGCAAGTGCAGAGTACTAGAGAAAGCAAGTCCCTGAGAT 1175  
QY 1176 CCGGCGGCGGCGAAGAGAGCAAGTGCAGAGTACTAGAGAAAGCAAGTCCCTGAGAT 1232  
DB 1081 GCTGTCCATGTGCGGCGGCGGCGGAGACAGAGTGTCAAGATCATGATGCGCTTCC 1140  
QY 1233 GCTGTCCATGTGCGGCGGCGGCGGAGACAGAGTGTCAAGATCATGATGCGCTTCC 1292  
DB 1141 AGAGCAGGAGAACTGAGAGAGAGATGCGGCGGAGTGTGAGTCCGCGCATCTGTA 1200  
QY 1293 AGAGCAGGAGAACTGAGAGAGAGATGCGGCGGAGTGTGAGTCCGCGCATCTGTA 1352  
DB 1201 CGAGCTGACCGAGCAGCATCAAGTTTCATCAACATGAACGCGCTTATGCGCGACCCAT 1260  
QY 1353 CGAGCTGACCGAGCAGCATCAAGTTTCATCAACATGAACGCGCTTATGCGCGACCCAT 1412  
DB 1261 GAAAGTGTACAAAGACCGCGGCTCATGAACATGAGTGAAGAGAGAGAGACTT 1320  
QY 1413 GAAAGTGTACAAAGACCGCGGCTCATGAACATGAGTGAAGAGAGAGAGACTT 1472  
DB 1321 CCGGAGAGAGTTCATGAGCAGATCCCAAGACTTTGAGCTGATGCGATCATCTTGAAG 1380  
QY 1473 CCGGAGAGAGTTCATGAGCAGATCCCAAGACTTTGAGCTGATGCGATCATCTTGAAG 1532  
DB 1381 GAAAGCAGTGTGAGTGTGCTCTTATTAATACTGAGTGAAGAGAGAGTGAAG 1440  
QY 1533 GAAAGCAGTGTGAGTGTGCTCTTATTAATACTGAGTGAAGAGAGAGTGAAG 1592  
DB 1441 GAGCTGTGAGAGAGAGTGTGCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAG 1500  
QY 1593 GAGCTGTGAGAGAGAGTGTGCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAG 1652  
DB 1501 GAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
QY 1653 GAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1712  
DB 1561 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
QY 1713 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1772  
DB 1621 CGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
QY 1773 CGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1832  
DB 1681 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
QY 1833 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1892



OY	1741	CCGATCAACCCGCTCAATGGCTAATGAGGCCCAACGACGAGAGGCGCATCACTCCCAAGCA	1800
Db	1893	CCGATCAACCCGCTCAATGGCTAATGAGGCCCAACGACGAGAGGCGCATCACTCCCAAGCA	1952
OY	1801	GAGCGCCGAGCTGGCTCCTCATGGACTGAATGAGAGTTCTGGCTTGGACAGAGAGAAAT	1860
Db	1953	GAGCGCCGAGCTGGCTCCTCATGGACTGAATGAGAGTTCTGGCTTGGACAGAGAGAAAT	2012
OY	1861	GGAAAACAGCCAAAGAAAGGTCTTCTTGGAAACAAGCCGCAACTGTGTGGCATCGCCGGAT	1920
Db	2013	GGAAAACAGCCAAAGAAAGGTCTTCTTGGAAACAAGCCGCAACTGTGTGGCATCGCCGGAT	2072
OY	1921	GSTGGGCTTCCAAGACTGTGTGCGCAGCTGTAAAGACTTCACTTCACTTCAAGAGAGCA	1980
Db	2073	GSTGGGCTTCCAAGACTGTGTGCGCAGCTGTAAAGACTTCACTTCACTTCAAGAGAGCA	2132
OY	1981	GAACTCGATGAGATCTTTCGACAGCAACAAGTGAAGATGGAGAAAGAGAGAAACGCGG	2040
Db	2133	GAACTCGATGAGATCTTTCGACAGCAACAAGTGAAGATGGAGAAAGAGAGAAACGCGG	2192
OY	2041	GAGGAAGAAAGAAAGAAAGCGCGCGCGCGCAGCAGAGAGGCTGATTCGCCCGCTGGT	2100
Db	2193	GAGGAAGAAAGAAAGAAAGCGCGCGCGCGCAGCAGAGAGGCTGATTCGCCCGCTGGT	2252
OY	2101	GAGAGATGAGAGATGAGAGCGGTGCGGCGTGAACGGAATGAGAGAGATGTGTAGAGAA	2160
Db	2253	GAGAGATGAGAGATGAGAGCGGTGCGGCGTGAACGGAATGAGAGAGATGTGTAGAGAA	2312
OY	2161	GGCTGAACCTTACATGCTCTGGGAATGAGAGTGGCCAGAGGGGAATGCATGGCCAGC	2220
Db	2313	GGCTGA-----AGC 2321	
OY	2221	CAGTGTCAACAAAGCTCAGACAACGAGAGATCCCTCTCCTCAACTGAGCGCGCAA	2280
Db	2332	CAGTGTCAACAAAGCTCAGACAACGAGAGATCCCTCTCCTCAACTGAGCGCGCAA	2381
OY	2281	GGAACAAGGGGAAGTGGGCCCCAAGCCCCCAGCCAACCTTGGGCGCGCAGCGGCAACCCC	2340
Db	2382	GGAACAAGGGGAAGTGGGCCCCAAGCCCCCAGCCAACCTTGGGCGCGCAGCGGCAACCCC	2441
OY	2341	AGGCCCAACCCCAACCCCAACGAGAGACATCCCGGGCCCCATTGAGCCCAACCCCGGCTC	2400
Db	2442	AGGCCCAACCCCAACCCCAACGAGAGACATCCCGGGCCCCATTGAGCTCAACCCCGGCTC	2501
OY	2401	TGAAGCAACCGGAAGCCCTCAAGCCCCCAACAGCAACCCCAACGCGCTCTGACCTCCGC	2460
Db	2502	TGAAGCAACCTTAAGCCCTTAAGCCCCCAACAGCAACCCCAACTTCTCTTACCTCTCC	2561
OY	2461	TGTGTGCCCAAGAGAGAGAGAGAGAGAGACCGACAGACGCCCCCAAGTGAAGAGAGG	2520
Db	2562	TGTGTGCCCAAGAGAGAGAGAGAGAGAGACCGACAGACGCCCCCAAGTGAAGAGAGG	2621
OY	2521	GAGAGAGCAAGACCCCCCGGCTGAGAGACTGGCACTGGCAACAGGGAAGGCCGAGAGAA	2580
Db	2622	GAGAGAGCAAGACCCCCCGGCTGAGAGACTGGCACTGGCAACAGGGAAGGCCGAGAGAA	2681
OY	2581	GCCGCTCAAGAGCGAGTGAACGAGAGGAAGCGAGAGAGGGGCGCGCAAGGCGCAAGAGCGC	2640
Db	2682	GCCGCTCAAGAGCGAGTGAACGAGAGGAAGCGAGAGAGGGGCGCGCAAGGCGCAAGAGCGC	2741
OY	2641	GAGGCGCTGATGAGCGCAAGCGCGAGAGGGGCGCTCAAGGCAAGAGAGAGAGAGGCGGAG	2700
Db	2742	GAGGCGCTGATGAGCGCAAGCGCGAGAGGGGCGCTCAAGGCAAGAGAGAGAGAGGCGGAG	2801
OY	2701	CGGCAAGGCGCAACTGCGCAAGAGCTGGGCGCGCCCCCAAGACAGCGACTTCAAGTGTAC	2760
Db	2802	CGGCAAGGCGCAACTGCGCAAGAGCTGGGCGCGCCCCCAAGACAGCGACTTCAAGTGTAC	2861
OY	2761	CTGCACTGACAGACGAGTGAAGTGAAGGCGGAGAGGCGGAGCAAGAAACGGGCTGTGTCCC	2820
Db	2862	CTGCACTGACAGACGAGTGAAGTGAAGGCGGAGAGGCGGAGCAAGAAACGGGCTGTGTCCC	2921
OY	2821	AAGGCCACGCTCTCAACCCCGAATGGCGAACCCCGGCGCAATGCTTCACTCCCAAGAGCC	2880

Db	2922	AAAGCCCA	GCCTCTC	CACCCG	AGCTGG	AGACCC	CCGGCC	CAATGCT	CAACCC	CAGAAC	CC	2988	
OY	2881	ACTGAC	CTGAAG	CAGCTGA	AGCAG	CGAGCT	GTGCAT	TCCGCC	CACTCA	AGCTCA	CCAA	2940	
Db	2982	ACTGGA	CTTGAAG	CACTGA	AGACAG	CGAGCT	GTGCAT	TCCGCC	CACTCA	AGCTCA	CCAA	3041	
OY	2941	AGTCCAT	GAGCCCC	CCGGAG	AGACAG	CGCTTCC	CAAGCC	AGTCCCC	CAGCCCC	AC	3000		
Db	3042	AGTCCAT	GAGCCCC	CCGGAG	AGACAG	CGCTTCC	CAAGCC	AGTCCCC	CAGCCCC	AC	3101		
OY	3001	GCCACCG	GAACCT	GCAGCC	GGAGAG	AGCCCT	CAAGAG	CGTTGAG	CAAGCC	CGCC	3066		
Db	3102	GCACCG	GAACCT	GCAGCC	GGAGAG	AGCCCT	CAAGAG	CGTTGAG	CAAGCC	CGCC	3167		
OY	3061	GGGCAAG	AGCAGAG	GC	CCGGAC	CCCCCG	CGACAG	AGAG-----			3100		
Db	3162	GGGCAAG	AGCAGAG	GC	CCGGAC	CCCCCG	CGACAG	AGAGGCGT	GTTCCT		3222		
OY	3101	----G	CTTGC	AGCCG	AGGCTC	AGAGAC	CCCTG	CTGTGA	CTTCCG		3156		
Db	3222	CCAGGCT	TTGCA	GGCCG	AGGCTC	AGAGAC	CCCTG	CTGTGA	CTTCCG		3283		
OY	3157	CCTG	CCCTT	CCCCCG	CGTGG	AGTATCA	AGGCTT	CCCCG	ATGCCCC	CGAACC	3216		
Db	3282	CCTG	CCCTT	CCCCCG	CGTGG	AGTATCA	AGGCTT	CCCCG	ATGCCCC	CGAACC	3341		
OY	3217	CTCAG	CCCTT	CTACG	CTTCA	CGCTG	CAACCA	CTGCGC	CTCCAT	GCATCA	CTGC	3276	
Db	3342	CTCAG	CCCTT	CTCTCA	CGCTTCA	CGTGGTCA	CCACTG	CCCTTGG	CTCTCA	GTGACAT	CTGC	3400	
OY	3277	CCGGCC	CGTCT	CGCG	CGCCCA	CCATCT	CCAA	CCCGCT	CCCTCAT	CTCTG	3336		
Db	3402	CCGGCC	CGTCT	CGCG	CGCCCA	CCATCT	CCAA	CCCGCT	CCCTCAT	CTCTG	3465		
OY	3337	CAAGCA	CCCA	CGCGTCT	CTGAG	AGCAAT	TGTTG	CCATTT	CTCCAG	GAATGT	CGGTCA	3396	
Db	3462	CAAGCA	CCCA	CGCGTCT	CTGAG	AGCAAT	TGTTG	CCATTT	CTCCAG	GAATGT	CGGTCA	3521	
OY	3397	GCTCA	CGTCC	CGTATCT	CAGAG	CAATG	CAAGCC	CCCGTGG	CCCTG	CACAT	TGGGCT	3456	
Db	3522	GCTCA	CGTCC	CGTATCT	CAGAG	CAATG	CAAGCC	CCCGTGG	CCCTG	CACAT	TGGGCT	3588	
OY	3457	GCCCC	TG	CCATG	AGCCCA	AAAGCT	TGCA	CCCTTCA	GC	CGAGT	GAAGCAG	3516	
Db	3582	GCCCC	TG	CCATG	AGCCCA	AAAGCT	TGCA	CCCTTCA	GC	CGAGT	GAAGCAG	3641	
OY	3517	GTCCCA	CGGGG	CAAG	GTGGG	CAAC	CGAGAG	CGCTGG	GGGTGCC	CAAGCC	CAGAGG	3576	
Db	3642	GTCCCA	CGGGG	CAAG	GTGGG	CAAC	CGAGAG	CGCTGG	GGGTGCC	CAAGCC	CAGAGG	3701	
OY	3577	GTCCG	TG	TGAAG	AGGACA	GTCTG	GGGTCA	GTCCGG	CGGAA	GCATCA	CAAAAGCAT	3638	
Db	3702	GTCCG	TG	TGAAG	AGGACA	GTCTG	GGGTCA	GTCCGG	CGGAA	GCATCA	CAAAAGCAT	3761	
OY	3637	TCCAG	CA	CAAGG	TG	CCCTCG	AGAG	CGGCAT	CAATAC	CGCGG	CTCATAC	3698	
Db	3762	TCCAG	CA	CAAGG	TG	CCCTCG	AGAG	CGGCAT	CAATAC	CGCGG	CTCATAC	3821	
OY	3697	CAGG	CA	CTGAG	CGTCT	GTGAC	AGGCA	CCATCA	CCAGAT	TCATCG	CGAGAG	3756	
Db	3822	CAGG	CA	CTGAG	CGTCT	GTGAG	AGGCA	CCATCA	CCAGAT	TCATCG	CGAGAG	3881	
OY	3757	GA	GTCCG	TG	TGAAG	AGGACA	GTCTG	GGGTCA	GTCCGG	CGGAA	GCATCA	3816	
Db	3882	GA	GTCCG	TG	TGAAG	AGGACA	GTCTG	GGGTCA	GTCCGG	CGGAA	GCATCA	3941	
OY	3817	CA	AAAGG	GC	CACTT	GTCTCT	TATG	AGGATG	CTGTGA	CCAGT	CTCTCA	3876	
Db	3942	CA	AAAGG	GC	CACTT	GTCTCT	TATG	AGGATG	CTGTGA	CCAGT	CTCTCA	4001	
OY	3877	GG	ACGG	CA	GAAG	CTTCA	AGAC	CCCCCAT	GACG	CGCG	CCCCCA	GGC	3936



Db 4002 GGA CGG CAG AAG CAG CTC A G A G C C C C C C A T G A G A C G C C C C C C C A A G C C C A C C T T A T G A 4061  
Qy 3937 CAT GAT G A G G G C C G C G T G G G C A G A G C C A T C T C T C A G C C A G C A T C G A A G T C T C A T G G G 3996  
Db 4062 CAT GAT G A G G G C C C G C T G G G C A G A G C C A T C T C T C A G C C A G A T C G A A G T C T C A T G G G 4121  
Qy 3997 C C G T C C A T C C C C G C G A G C C A C A G C C C C C A C C T C A A A G A G C A G C A C A C A T C C G 4056  
Db 4122 C C G T C C A T C C C C G C G A G C C A C A G C C C C C A C C T C A A A G A G C A C C A C A T C C G 4181  
Qy 4057 C G G G T C C A T C A C A C A A G G A T T C C T C G T C T A C G T G A G G C A C A G A G A C T T A C T T G G 4116  
Db 4182 C G G G T C C A T C A C A C A A G G A T C T C G T C T A C G T G A G G C A C A G A G A C T T A C T T G G 4241  
Qy 4117 T C G G A G G C C A A G C T C T T A A C G G G A G G C A G C C T C G C C C C A C G C C C T C A C G G A 4176  
Db 4242 T C G G A G G C C A A G C T C T T A A C G G G A G G C A G C C T C G C C C C A C G C C C T C A C G G A 4301  
Qy 4177 C C T G A C C G A G G C T T A C A A G C G C A G G C C C T G G G C C C C T T G A A G C T T G A A G C C G G C C A T G A 4236  
Db 4302 C C T G A C C G A G G C C T T A C A A G C G C A G G C C C T G G G C C C C T T G A A G C T T G A A G C C G G C C A T G A 4361  
Qy 4237 G G G C C T G T G C C A C A G G T G A A G A G G C G G C C G C T C A T C A T G A G A T C C G C G C G A G G A 4296  
Db 4362 G G G C C T G T G C C A C A G G T G A A G A G G C G G C C G C T C A T C A T G A G A T C C C G C G A G G A 4421  
Qy 4297 G C T G C G G C A C A G C C C G A G C T C C C C T G G C C C C G C G C C C T C A A G A G G G C T C C A T C A C 4356  
Db 4422 G C T G C G G C A C A G C C C G A G C T C C C C T G G C C C C G C G C C C T C A A G A G G G C T T C A T C A C 4481  
Qy 4357 G C A G G G C A C C C C G C T C A A G T A C G A C A C C G G C C G T C A C C A C T G G C T C A A A A G C A C G A 4416  
Db 4482 G A G A G G C A C C C C G C T C A A G T A C G A C A C C G G C C G T C A C C A C T G G C T C A A A A G C A C G A 4541  
Qy 4417 C G T A G C T C C C T C A T C G G A G C C C C G C G A C G T T C C C A C C C C T G C A C C C C T G A T G T 4476  
Db 4542 C G T A G C T C C C T C A T C G G A G C C C C G C G A C G T T C C C A C C C C T G A T G T 4601  
Qy 4477 G A T G C C G A C G C C C G G G C A C T G A A C G T G C T G A C G A G A G C C T G A A G C C G A G C 4536  
Db 4602 G A T G C C G A C G C C C G G G C A C T G A A C G T G C T G A C G A G A G A G C C T T A A A G C C G A G C 4661  
Qy 4537 A G G G A C C G C A G C A G C T C G G G G G C T C A T T G C G C G C G C C C G G T C A T T G T G C T T G A 4596  
Db 4662 A G G A C C G C A G C A G C T C G G G G G C T C A T T G C G C G C C C C G G T C A T T G T G C T T G A 4721  
Qy 4597 G C T G G G T A G C C C G C G A G A G C C C C T T G A C T A T G A G A C A C G G G G C A C C C T T T G C C G G 4656  
Db 4722 G C T G G G T A G C C C G C G A G A G C C C C T T G A C T A T G A G A C A C G G G G C A C C C T T T G C C G G 4781  
Qy 4657 C C A C C T C C C A C A G A G T T G C C C C G T G A C A T G C G G G A G C C A C G C C G C C T G A G A G G G 4716  
Db 4782 C C A C C T C C C A C A G A G T T G C C C C G T G A C A T G C G G G A G C C C A G C C G C C T G A G A G G G 4841  
Qy 4717 C A G C C T T T C G T C A G C A A G A C A T C C C A G A C C G A A A G C T A G C T G A C G C C T G T G A G A T 4776  
Db 4842 C A G C C T T T C G T C A G A G A G C A T C C C A G A C C G A A A G C T A G C T G T G A G A T 4901  
Qy 4777 C G C C A A G T C C C C G C A C A G A C C G T G C C G A G A C C A C C C A C C C C A T T C G C C C T A T G A 4836  
Db 4902 C C C C A A G T C C C C G C A C A G A C C G T G C C C G A G A C C A C C A C A C C C C A T C T C G C C C T A T A 4961  
Qy 4837 G C A C C T G C T T G G G G G T G A G T G G G T G G A C C T G T A T G C A G C A C A T C C C C T G G G C C T T 4896  
Db 4962 G C A C C T G C T T G G G G G T G A G T G G G A C C T G T A T G C A G C A C A T C C C C T G G C C T T 5021  
Qy 4897 C G A C C C C A C C T C A T A C C C C G C G A T C C C T T G A C G A C G A C C G C T C C T A C T A C T G C C 4956  
Db 5022 C G A C C C C A C C T C A T A C C C C G G G A T C C C T C T G A C G A C G C G C T T A C T A C T A C T G C C 5081  
Qy 4957 C C G A C A C C T G C C C C C A C C C C A C C T A C C C G A C C T G T A C C A C C C T A C C T A C C G G G 5016  
Db 5082 C C G A C A C C T G C C C C C A C C C C A C C C C A C C T A C C C G A C C T G T A C C A C C C T A C C T A T C C G G G 5141

Qy 5017 C T A C C C C G A C A C G C G G C G T G A G A C C G G C A G C A T C A T C A A T G A C T A C A T C A C C T C 5076  
Db 5142 C T A C C C C A C A C G G G G G C T G A G A A C C G G C A G C A T C A T C A A T G A C T A C A T C A C C T C 5201  
Qy 5077 G C A G A G A T G C A C C A C A C A C G G C C A C C G C C A T G G C C C A G C A G C T G A T A T C T A G G G G 5136  
Db 5202 G C A G A G A T G C A C C A C A C A C G G C C A C C G C C A T G G C C C A G C A G C T G A T A T C T A G G G G 5261  
Qy 5137 C C T C T G C C C C G A G T C C T G C T G C A C T C A A C T A C G C T G G G G T C C C C A G C A T C A T 5196  
Db 5262 C C T C T G C C C C G A G T C C T G C T G C A C T A C G T G C G G G T C C C C A G C A T C A T 5321  
Qy 5197 C G A C C T G T C C C A A G T G C A C A C T G C C T G T G C C C C G A C A C A G G A C C C C A G C 5256  
Db 5322 C G A C C T G T C C C A A G T G C C A C C T G C C T G T G C C C C G A C A C A G G A C C C C A G C 5381  
Qy 5257 C A C C G C A T G A C C G C C T T T G C C T A C C T C C C A C C G G C C C A G C C C T T C A G A G C C G C C A 5316  
Db 5382 C A C C G C A T G A C C G C C T T T G C C T A C C C C C A C C G G C C C A G C C C T T C A G A G C C G C C A 5441  
Qy 5317 C A G C A C T C C C C A C T C C C C C A G A G T C C A C A C A C T T T G A C A A A C C A A C C A C C A C G T C 5376  
Db 5442 C A G C A C T C C C A C T C C C C A G A G T C C A C A C T T T G A C A A A C C A A C C A C C A C G T C 5501  
Qy 5377 C T C G T C C G A G C G G A G C G A C C G G A T C G A G C G G G A C C G G A T C G G A G C G G A A A 5436  
Db 5502 C T C G T C C G A G C G G A G C G A G C C G G A T C G A G C G G G A C C G G A T C G G A G C G G A A A 5561  
Qy 5437 G T C A T T C C T C A C G T C A C A C A C A C G T G A G A C G C A C C C A C T T T G A A C C T G T A C A G A 5496  
Db 5562 G T C A T T C C T C A C G T C A C A C A C A C G T G A G A C C C A C C A C T T T G A A C C T G T A C A G A 5621  
Qy 5497 G C A G A G C G C G G C A C A G C G G C A G C A G C G G G G T G G G G G C A G A C A G C C G C C G C G 5556  
Db 5622 G C A G A G C G C G C A C A G C G G C A G A G C G G G G T G G G G G C A G C A G C C G C C G C G 5681  
Qy 5557 C T C C C A C T C C C A T G C C A C A C A C T G C C C A T C T C C C C T C G A C C C A G A T G C C T C C A 5616  
Db 5682 C T C C C A C T C C C A T G C C A C A C A C T G C C C A C T C C C C T C G A C C C A G A T G C C T C C A 5741  
Qy 5617 G C A G A G C C C A G T G C T T C A C A C A C A G C G A T G A A G G T A T C A T C A C C C G C T G T G A G C C 5676  
Db 5742 G C A G A G C C C A G T G C T T C A C A C A C A G C G A T G A A G G T A T C A T C A C C C G C T G T G A G C C 5801  
Qy 5677 C A G C A G C C C A C G T C C T G A G G T C C A C C T C C A C C C T C A C C C G T T C G C C A G C T G C C A C 5736  
Db 5802 C A G C A G C C C A C G T C T G A G G T C C A C C T C A C C T C C A C C C G T T C G C C A G C T G C C A C 5861  
Qy 5737 A T T C C A C C T G C C A C C A C T G C C C A C T G G G C G A C C C T C A T G A G G G T C T A C C T A C C C T 5796  
Db 5862 A T T C C A C C T G C C A C C A C T G C C C A C T G G G G G A C C C T C A T G A G G G T C T A C C T A C C C T 5921  
Qy 5797 C A T G A G C C C G T C T T G C T G C C C A G A G G C C C C G G G T C C C C G C C A G A G C G C C C G 5856  
Db 5922 C A T G A G C C C G T C T T G C T G C C C A G A G G G C C C C G G G T C C C C G C C A G A G C G C C C G 5981  
Qy 5857 A G C A G A C C C G C C A T G C C T T C T G C C A G A G C C C A G C C C G C C T C G G G T G A A G C C C G C 5916  
Db 5982 A G C A G A C C C G C C A T G C C T T C T G C C A G A G C C C C A G C C C G C C T C G G G C T G A A G C C C G C 6041  
Qy 5917 C T C T C C C C A G C A A G G C T C G A A C C C C G G C C C C T A G T A G C T C C T G T C T G C C A C G C 5976  
Db 6042 C T C C T C C C C A A A G G G C T C G A A C C C C G G C C C C T A G T A G C T C C T G T C T G G C C A G C 6101  
Qy 5977 C A C A T G C C G C C A C C C T T G C G A A A C C T G C A C C T C A C A G C C A G C C C G A C C C G C C 6036  
Db 6102 C A C A T G C C C C A C C C C T T G C G A A A C C T G C A C C T C A C A G C C A G C C C G A C C C G C C 6161  
Qy 6037 G G C G C A C C T G C T G G C C T C G A A C C C G A C C G G A A A A G A C T C A A A G T A A A C C C T T T C 6096  
Db 6162 G G C G C A C C T G C T G G C C T C G A C C C G A C C G G A A A A G A C T C A A A G T A A A C C C T T T C 6221

QY 6097 CATCCGGAAGCTGGAACTCGTTCTCTGGGTTACACGGGACAGCTTACAGCCCGGAAGG 6156  
DB 6222 CATCCGGAAGCTGGAACTCGTTCTCTGGGTTACACGGGACAGCTTACAGCCCGGAAGG 6281  
QY 6157 GGTGGAGCCCGTCAAGCCCTGTGAGCTCAACCAGTGTGACCACAGAACAGGGGCTCCCA 6216  
DB 6282 GGTGGAGCCCGTCAAGCCCTGTGAGCTCAACCAGTGTGACCACAGAACAGGGGCTCCCA 6341  
QY 6217 GCACTTGAAGAGCTGACAGAGAGCACTTGAAGGGAGCTGGCCCAAGAGCCAGG 6276  
DB 6342 GCACTTGAAGAGCTGACAGAGAGCACTTGAAGGGAGCTGGCCCAAGAGCCAGG 6401  
QY 6277 CCCCGTGAAGCTTGGCGGGAGGCGCCCACTTCCACACTTGGCGCGCTGCTTGAAG 6336  
DB 6402 CCCCGTGAAGCTTGGCGGGAGGCGCCCACTTCCACACTTGGCGCGCTGCTTGAAG 6461  
QY 6337 CCAGCCCTGTCCAGGCGCGCTGCTCAAGACCGGCGCAAGGGGTCAAGGTCAACAGCGGGT 6396  
DB 6462 CCAGCCCTGTCCAGGCGCGCTGCTCAAGACCGGCGCAAGGGGTCAAGGTCAACAGCGGGT 6521  
QY 6397 GGTCACTCTGGCCCAAGCACTAGTAGGTCACTACAGAGACTAGACCCGGACCAACC 6456  
DB 6522 GGTCACTCTGGCCCAAGCACTAGTAGGTCACTACAGAGACTAGACCCGGACCAACC 6581  
QY 6457 ACAGCAGCTCAGGCAACCCCTGGCGCGCCCTCTACTCTCTTCCCTGGGGCAGCTGCC 6516  
DB 6582 ACAGCAGCTCAGGCAACCCCTGGCGCGCCCTCTACTCTCTTCCCTGGGGCAGCTGCC 6641  
QY 6517 CGTCTTGAAGCTTCCGCGCGCCCAAGTAGTACTTCTTCCGCGCCCGGACCAATGGTGC 6576  
DB 6642 CGTCTTGAAGCTTCCGCGCGCCCAAGTAGTACTTCTTCCGCGCCCGGACCAATGGTGC 6701  
QY 6577 CCGGCGCGGTGGCTCCCGCAAGGAGGGGCAAGGGTCTCCAGAGCCAAACAAGAC 6636  
DB 6702 CCGGCGCGGTGGCTCCCGCAAGGAGGGGCAAGGGTCTCCAGAGCCAAACAAGAC 6761  
QY 6637 GTTCGCTTGGGTGTGTGAGGACGGTATTGAACCTGTGTCCCAACCGGAGGGCAATGAC 6696  
DB 6762 GTTCGCTTGGGTGTGTGAGGACGGTATTGAACCTGTGTCCCAACCGGAGGGCAATGAC 6821  
QY 6697 GGAAGCCAGGGCACTCCCGAGTGTGTGTAACCGCTGTGTACCGGGATGGGGACAAGAC 6756  
DB 6822 GGAAGCCAGGGCACTCCCGAGTGTGTGTAACCGCTGTGTACCGGGATGGGGACAAGAC 6881  
QY 6757 GGAAGCCAGGAGATGGGCTCCAGTCTCCAGGCAACACGAGCCAGCCGCAAGCTTCTT 6816  
DB 6882 GGAAGCCAGGAGATGGGCTCCAGTCTCCAGGCAACACGAGCCAGCCGCAAGCTTCTT 6941  
QY 6817 CAGCAGCTGACCGAGAGCACTCCGCAATGTCAAGTCCAAAGAACAGATCAACA 6876  
DB 6942 CAGCAGCTGACCGAGAGCACTCCGCAATGTGTCAAGTCCAAAGAACAGATCAACA 7001  
QY 6877 GAAAGCTGAACCCCAACACCGGAATGAGCTTGAATACATATCAGCCAGCTTGGACGGA 6936  
DB 7002 GAAAGCTGAACCCCAACACCGGAATGAGCTTGAATATCAAGCTTGGACGGA 7061  
QY 6937 GATCTTCAATATGCCCCGCACTACCCGGAACAGGCTTATGACCTTATAGAAAGCCAGCGGT 6996  
DB 7062 GATCTTCAATATGCCCCGCACTACCCGGAACAGGCTTATGACCTTATAGAAAGCCAGCGGT 7121  
QY 6997 GAGAGAACATGCGACGCAACATGAGGGCTGAGAGCCATATATTAAGAAAGCACTCAATGGG 7056  
DB 7122 GAGAGAACATGCGACGCAACATGAGGGCTGAGAGCCATATTAAGAAAGCACTCAATGGG 7181  
QY 7057 TAAATATGACAGTGGGAAGAGTCCCGCGCTCAGCGCCAAATGCTTTTAAACCTCTGA 7116  
DB 7182 TAAATATGACAGTGGGAAGAGTCCCGCGCTCAGCGCCAAATGCTTTTAAACCTCTGA 7241  
QY 7117 TGGCAGTGGCAAGCTTGGCGCGCTGTATGCCCCAATAACCGCTGTGAAGAGAGGTACCA 7176  
DB 7242 TGGCAGTGGCAAGCTTGGCGCGCTGTATGCCCCAATAACCGCTGTGAAGAGAGGTACCA 7301  
QY 7177 CACACTCACTCGCAGGTGGCGGGGGAAGGCCAAGGTCTGTGAGACCCAGCAGCCG 7236

DB 7302 CACACTCACTCGCAGGTGGCGGGGAAGGCCAAGGTCTGTGAGACCCAGCAGCCG 7361  
QY 7237 AAAAGCAAGTCCCGGACCCCGGGCTGGCATTTGGGGAACCGGCAACCTCTGTCTTC 7296  
DB 7362 AAAAGCAAGTCCCGGACCCCGGGCTGGCATTTGGGGAACCGGCAACCTCTGTCTTC 7421  
QY 7297 AGTGCACTGGAGGAGAGCTGCAACCGCGGAGCGCGCTCAACCAACCGCGTGTGGAGGA 7356  
DB 7422 AGTGCACTGGAGGAGAGCTGCAACCGCGGAGCGCGCTCAACCAACCGCGTGTGGAGGA 7481  
QY 7357 CAGGCCCTGTCCGAGAGTTTCCAGCCATTCCTTCAACCCCTGATATGAGGCTGCA 7416  
DB 7482 CAGGCCCTGTCCGAGAGTTTCCAGCCATTCCTTCAACCCCTGATATGAGGCTGCA 7541  
QY 7417 GGGGGGTGATAGGCTTCCCAACCGGGCTCCCGAGGAGAGCGGGCCCTCGC 7476  
DB 7542 GGGGGGTGATAGGCTTCCCAACCGGGCTCCCGAGGAGAGCGGGCCCTCGC 7601  
QY 7477 TGGCCCCCAACAGCCTGGGACAGAGAGCCCAAGCCACTGTCTGTGCAATGACAGAC 7536  
DB 7602 TGGCCCCCAACAGCCTGGGAGAGAGAGCCCAAGCCACTGTCTGTGCAATGACAGAC 7661  
QY 7537 ACTTCCGACAGCGAGTGACTAGAAACAGGCGGGGGGGGGCGGGCGGTGERSPERLUG 7596  
DB 7662 ACTTCCGACAGCGAGTGACTAGAAACAGGCGGGGGGGGGCGGGCGG-----TG 7711  
QY 7597 TCAAGTCCGAGGAGCAAGGAAAGGCGCTGACAGAGGGGGGGGCTGCGACTCCCC 7656  
DB 7712 TCAAGTCCGAGGAGCAAGGAAAGGCGCTGACAGAGGGGGGGGCTGCGACTCCCC 7771  
QY 7657 AACCAAGAAAGAGCCCTTGAAGCTCGCTGCGCTCATTCATCTGTCTGCAAGAGCG 7716  
DB 7772 AACCAAGAAAGAGCCCTTGAAGCTCTTAACTTAACTCCGCGCGGGCTGGGCTGTACAGA 7831  
QY 7717 GCATCTTGCCTGTCTTAAAGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 7776  
DB 7832 GCATCTTGCCTGTCTTAAAGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 7891  
QY 7777 CCTTACTAGGGGAGTCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 7836  
DB 7892 CCTTACTAGGGGAGTCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 7951  
QY 7837 GCAAGGCAAGGCTGTGGGAGGCAACAGAGGCGGCAAGGGGGGCAAGGACCCAAAGCAG 7896  
DB 7952 GCAAGGCAAGGCTGTGGGAGGCAACAGAGGCGGCAAGGGGGGCAAGGACCCAAAGCAG 8011  
QY 7897 GATGACCAAGCACTTCCAGGCACTGCTCCCGAATGCAATTGGAACCAAGTCTTAA 7956  
DB 8012 GATGACCAAGCACTTCCAGGCACTGCTCCCGAATGCAATTGGAACCAAGTCTTAA 8071  
QY 7957 CTGAGCTCGAGGCCCCCGGCGCTCCCTCGGCTCCCATTCGCGCTTGAAGCAG 8016  
DB 8072 CTGAGCTCGAGGCCCCCGGCGCTCCCTCGGCTCCCATTCGCGCTTGAAGCAG 8131  
QY 8017 ATGGAACGAGGCGCTGTCCAGGCGCGGAGGCGTGGTCCGACCAAGACTGCCCCA 8076  
DB 8132 ATGGAACGAGGCGCTGTCCAGGCGCGGAGGCGTGGTCCGACCAAGACTGCCCCA 8191  
QY 8077 GCCAAGCAATTTGCTGGAACCAAGTCAAGGCAAGTGGGCGGACAAAGGGCCAGGTGCG 8136  
DB 8192 GCCAAGCAATTTGCTGGAACCAAGTCAAGGCAAGTGGGCGGACAAAGGGCCAGGTGCG 8251  
QY 8137 GCGTGGGGGAACGAGTCTCCAGAGACTGGAATGTTTTTTTCAACATCTGTCGCGAG 8196  
DB 8252 GCGTGGGGGAACGAGTCTCCAGAGACTGGAATGTTTTTTTCAACATCTGTCGCGAG 8311  
QY 8197 CGGTGGGAAGAAAGCAGATGTAATGATGTGTTTAAAGGATATATTTTGAATAC 8256  
DB 8312 CGGTGGGAAGAAAGCAGATGTAATGATGTGTTTAAAGGATATATTTTGAATAC 8371  
QY 8257 CTTCAATGAATTAATTCAAGATGTTTTTACAAAGAAAGACTTAAACCAAGTATTACTGCTCG 8316

Db 8372 CTTCAATGAATTAATTCAGANGTTTTAGCAAGAGGACCTTACCAGTATTACTGCTGC 8431  
QY 8317 TGTGCTTTGATCTCTGCTTACCGTTCAAGAGCGCTGTGCAAGCCGACAGTGCATGCC 8376  
Db 8432 TGTGCTTTGATCTCTGCTTACCGTTCAAGAGCGCTGTGCAAGCCGACAGTGCATGCC 8491  
QY 8377 CATCATCGCAGAGCAAGAGGGGGCGGGGACGTCTGTCAAGCCCGCTGTCTCTCCCTC 8436  
Db 8492 CATCATCGCAGAGCAAGAGGGGGCGGGGACGTCTGTCAAGCCCGCTGTCTCTCCCTC 8551  
QY 8437 CCTCCCTTCCTTGGGCAAGAAATTAATTCAGATGCGTATTCTGTGGCCGCCATTGGCGAGGG 8496  
Db 8552 CTTCCCTTCCTTGGGCAAGAAATTAATTCAGATGCGTATTCTGTGGCCGCCATTGGCGAGGG 8611  
QY 8497 TGTGCTATTCTGTCAATTACACAGTCGCTTAAATTAAGCAATTAATCTCCAAA 8556  
Db 8612 TGTGCTATTCTGTCAATTACACAGTCGCTTAAATTAAGCAATTAATCTCCAAA 8671  
QY 8557 AAAAAAAAAAAAAA 8571  
Db 8672 AAAAAAAAAAAAAA 8686

## RESULT 5

US-10-174-014-4  
; Sequence 4, Application US/10174014  
; Publication No. US20040005292A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Susan M. Freier  
; APPLICANT: Kenneth W. Dobie  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION  
; FILE REFERENCE: PTS-0012  
; CURRENT APPLICATION NUMBER: US/10/174,014  
; CURRENT FILING DATE: 2002-06-17  
; NUMBER OF SEQ ID NOS: 73  
; SEQ ID NO 4  
; LENGTH: 8686  
; TYPE: DNA  
; ORGANISM: H. sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (157)...(7680)  
US-10-174-014-4

Query Match 97.2%; Score 8324.6; DB 6; Length 8686;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 8472; Conservative 1; Mismatches 34; Indels 88; Gaps 4;

QY 1 CATGTCGAGGCTCCACACAGCTTGTGCAAGAGCGGCACTGAGCCCGCTACCC 60  
Db 156 CATGTCGAGGCTCCACACAGCTTGTGCAAGAGCGGCACTGAGCCCGCTACCC 215  
QY 61 GCCCCACAGCCTTTCTTACCCAGTGCAGATCGCCCGAGCGCACAGGAGTGGCTCTCT 120  
Db 216 GCCCCACAGCCTTTCTTACCCAGTGCAGATCGCCCGAGCGCACAGGAGTGGCTCTCT 275  
QY 121 GAGATACCAAGCACTCCCGGCACTATGCTCCCACTGTGGCGGGCTCATATCA 180  
Db 276 GAGATACCAAGCACTCCCGGCACTATGCTCCCACTGTGGCGGGCTCATATCA 335  
QY 181 GCCCCAGCGGCGAGAGGCTCTCCGTGCTGTAGATTCAAGCCCGGAAATGAACGTTCCA 240  
Db 336 GCCCCAGCGGCGAGAGGCTCTCCGTGCTGTAGATTCAAGCCCGGAAATGAACGTTCCA 395  
QY 241 GAGAGTCCACCTGCGAGCGAGTCCCATCATACCTGCCCGAGCTGGGGAATCAGAGAT 300  
Db 396 GAGAGTCCACCTGCGAGCGAGTCCCATCATACCTGCCCGAGCTGGGGAATCAGAGAT 455  
QY 301 GAGAGTTCATTGAAGAGAGCGCCCTGCGCTTAAAGCTGCTGACCCCTGCTGACCC 360  
Db 456 GAGAGTTCATTGAAGAGAGCGCCCTGCGCTTAAAGCTGCTGACCCCTGCTGACCC 515

QY 361 GTCAACCTCTGTGGCCACAGGGCCAGCCTTGGGGATCTGAAGACTCTACCAAGAGCCGTAG 420  
Db 516 GTCAACCTCTGTGGCCACAGGGCCAGCCTTGGGGATCTGAAGACTCTACCAAGAGCCGTAG 575  
QY 421 CTTGACGGGCAAGCTGTGAACCGGTGTCTCCCGCCAGCCCCCGGCACTGTACCCCTGAGCT 480  
Db 576 CTTGACGGGCAAGCTGTGAACCGGTGTCTCCCGCCAGCCCCCGGCACTGTACCCCTGAGCT 635  
QY 481 GAGAGTGTGCGCGACCGGCTGTCCAAAGAGAGCTGATCCAGAACATGAGCCGCTGGA 540  
Db 636 GAGAGTGTGCGCGACCGGCTGTCCAAAGAGAGCTGATCCAGAACATGAGCCGCTGGA 695  
QY 541 CCGAGAGATCAACCATGTGAGAGCAGAGATCTTAAGCTGAAGAAAGCAGCAACAGCT 600  
Db 696 CCGAGAGATCAACCATGTGAGAGCAGAGATCTTAAGCTGAAGAAAGCAGCAACAGCT 755  
QY 601 GAGAGAGAGGCTGTGCAGACCGGCCGAGCCGTAAGAGCCCGGTTCACCGCCGCATGGA 660  
Db 756 GAGAGAGAGGCTGTGCAGACCGGCCGAGCCGGTTCACCGCCGCATGGA 815  
QY 661 GTCAAGAGCAGCAGCCTGTGTGAGATCATCTACAGACGAGAACCGAGAAAGCTGAAGC 720  
Db 816 GTCAAGAGCAGCAGCCTGTGTGAGATCATCTACAGACGAGAACCGAGAAAGCTGAAGC 875  
QY 721 TGCAATCGAATTTTGAAAGGCTTGGGGCCCAAGGTGAGCTGCCGTGTACCAACGACC 780  
Db 876 TGCAATCGAATTTTGAAAGGCTTGGGGCCCAAGGTGAGCTGCCGTGTACCAACGACC 935  
QY 781 CTCGACACCCCGGCAAGTATCATGAGAAATCAAAATTAACCAAGCGAGTCCGAGAAAGCT 840  
Db 936 CTCGACACCCCGGCAAGTATCATGAGAAATCAAAATTAACCAAGCGAGTCCGAGAAAGCT 995  
QY 841 AATCTTACTTCAAGAGAGGAATCAGCTCGAGAAACATGAGAGCAGAAATTTGTGCA 900  
Db 996 AATCTTACTTCAAGAGAGGAATCAGCTCGAGAAACATGAGAGCAGAAATTTGTGCA 1055  
QY 901 GCGCTATGACCAAGCTCATGAGAGCCTTGGAAAAAAGGTGAGCGCATGAAAACAACCC 960  
Db 1056 GCGCTATGACCAAGCTCATGAGAGCCTTGGAAAAAAGGTGAGCGCATGAAAACAACCC 1115  
QY 961 GCGCGGCGGGCCAAAGAGAGCAAGTGCAGAGTACTACGAAAGCATTTCCCTGAGAT 1020  
Db 1116 CCGGCGGCGGGCCAAAGAGAGCAAGTGCAGAGTACTACGAAAGCATTTCCCTGAGAT 1175  
QY 1021 CCGGAGAGAGCGGAGCTGTGAGAGCGCATGAGAGCAGGCTGGGCGAGCGGGGAGTGG 1080  
Db 1176 CCGGAGAGAGCGGAGCTGTGAGAGCGCATGAGAGCAGGCTGGGCGAGCGGGGAGTGG 1232  
QY 1081 GCTGTCCATGTTCGCGCGCCCGAGGAGCAGAGGTGTCAAGATCATGATGCGCTCTC 1140  
Db 1233 GCTGTCCATGTTCGCGCGCCCGAGGAGCAGAGGTGTCAAGATCATGATGCGCTCTC 1292  
QY 1141 AGAGCAGAGAACTGTGAAGAAAGCAGATGCCAGCTGGCCGTGATCCCGCCATCTCTGA 1200  
Db 1293 AGAGCAGAGAACTGTGAAGAAAGCAGATGCCAGCTGGCCGTGATCCCGCCATCTCTGA 1352  
QY 1201 CGAGCTGACCAAGAGCGGCATCAAGTTTCATCAACATGAACGGGCTTATGGCCGACCCAT 1260  
Db 1353 CGAGCTGACCAAGAGCGGCATCAAGTTTCATCAACATGAACGGGCTTATGGCCGACCCAT 1412  
QY 1261 GAAAGTGTAAAGAACCGCCAGCTCATGAACATGTGAGTGAAGCAGAGAAAGAGACCTT 1320  
Db 1413 GAAAGTGTAAAGAACCGCCAGCTCATGAACATGTGAGTGAAGCAGAGAAAGAGACCTT 1472  
QY 1321 CCGGAGAAAGTTTCATGACAGATCCCAAGAACTTTGGCTGATCGATATTTCTTGAAGAG 1380  
Db 1473 CCGGAGAAAGTTTCATGACAGATCCCAAGAACTTTGGCTGATCGATATTTCTTGAAGAG 1532  
QY 1381 GAAAGCAGTGTGAGTGTGCTCTATTACTACCTGAAGAAAGTGAAGCATATTA 1440  
Db 1533 GAAAGCAGTGTGAGTGTGCTCTATTACTGAAGAAAGTGAAGCATATTA 1592

[illegible]

Db	2622	GGAGGAGCAGAAAGCCCCCGCGCGCTGAGAGACTGGCAGTGGACACAGGAAAGCCGAGGA	2681
Oy	2581	GCCCCCTCAAGAGCGAGTGCACGAGAGAAAGCCGAGAGAGGGGCGCGGCAAGGGCAAGACGC	2640
Db	2682	GCCCCCTCAAGAGCGAGTGCACGAGAGAAAGCCGAGAGAGGGGCGCGGCAAGGGCAAGACGC	2741
Oy	2641	GAGAGCGCTGAGAGGCACGAGCCGAGAGGGGCGCTCAAGCAGAGAAAGAGAGGGCGGAG	2700
Db	2742	GAGAGCGCTGAGAGGCACGAGCCGAGAGGGGCGCTCAAGCAGAGAAAGAGAGGGCGGAG	2801
Oy	2701	CGGAGAGGGGCGACACTGCGAAGACTCGGGGGGCCCCCAAGACAGGGCACTCAGTGCAC	2766
Db	2802	CGGAGAGGGGCGACACAGCCAAAGACTCGGGGGGCCCCCAAGACAGGGCACTCAGTGCAC	2861
Oy	2761	CTGCAGTGCACACAGAGTGAATGAGAGCGAGAGGGCGGCGACAAGAACCGGCTGTCTCCCC	2820
Db	2862	CTGCAGTGCACACAGAGTGAATGAGAGCGAGAGGGCGGCGACAAGAACCGGCTGTCTCCCC	2921
Oy	2821	AAGGCCACAGCCTCTCACCCCGGACTGGCGACCCCGGGCCAAATGCTTCAACCAGAAAGCC	2880
Db	2922	AAGGCCACAGCCTCTCTCACCCCGGACTGGCGACCCCGGGCCAAATGCTTCAACCAGAAAGCC	2981
Oy	2881	ACTGGACTTGAAGCACTGAAGCAGCGGAGCGGCTGGCCATCCCCCCCATCCAGGTCAACAA	2940
Db	2982	ACTGGACTTGAAGCACTGAAGCAGCGGAGCGGCTGGCCATCCCCCCCATCCAGGTCAACAA	3041
Oy	2941	AGTCATATAGCCCCCGGGGAGAGCGGAGCTCCGACCAAGCAAGCACTCCCCAGGCCCAAC	3000
Db	3042	AGTCATATAGCCCCCGGGGAGAGCGGAGCTCCGACCAAGCAAGCACTCCCCAGGCCCAAC	3101
Oy	3001	GCCACCGCAAAACCTGCAGCGCGGAGAGCGACGCCCTTCAGCAGCGCTGGCAGCAGCCCCG	3060
Db	3102	GCCACCGCAAAACCTGCAGCGCGGAGAGCGACGCCCTTCAGCAGCGCTGGCAGCAGCCCCG	3161
Oy	3061	GGGCAAGAGCAGAGAGCCCGGCACCCCCGCCGACAGAG-----31000	
Db	3162	GGGCAAGAGCAGAGAGCCCGGCACCCCCGCCGACAGAGAGCAGAGAAAGCCTGTGTTCTT	3221
Oy	3101	----GCCTTGGCACCCGAGGCCGAGAAAGCTGGCGGGACCCCCCTTGCCTGGACCTTCCGG	3156
Db	3222	CCGAGCCTTTCGCAACCCGAGAGCCCGAGAAAGCTGCGGGACCCCCCTTGCCTGGACCTTCCGG	3281
Oy	3157	CCTGCGCCTTCCCGGTGCGCCCCCGCTGAGAGTATCAAGGCTCTCCCGCATGCCCCCGGACCC	3216
Db	3282	CCTGCGCCTTCCCGGTGCGCCCCCGCTGAGAGTATCAAGGCTCTCCCGCATGCCCCCGGACCC	3341
Oy	3217	CTCAGCCTTCTCTACGCTCCACTGTGTCAACCACTGCGCCCTTGGGGCTTCACATGACACTGC	3276
Db	3342	CTCAGCCTTCTCTACGCTCCACTGTGTCAACCACTGCGCCCTTGGGGCTTCACATGACACTGC	3401
Oy	3277	CCGGCCCGTCTGCGCGGGGCCACCCACATTTCCAAACCCGCTCTCCCTCATTTCTCTCTCG	3336
Db	3402	CCGGCCCGTCTGCGCGGGGCCACCCACATTTCCAAACCCGCTCTCCCTCATTTCTCTCTCG	3461
Oy	3337	CAAGCACCACCGAGCTCTCGAGAGGAGCAATATGAGGCCATCTCCAGAGAAATGTGCTCA	3396
Db	3462	CAAGCACCACCGAGCTCTCGAGAGGAGCAATATGAGGCCATCTCCAGAGAAATGTGCTCA	3521
Oy	3397	GCTCCAGTCCCGTACTCAGAGCATGCCAAAGGCCCGGGTGGGCCCTGTGCACATGGGGCT	3456
Db	3522	GCTCCAGTCCCGTACTCAGAGCATGCCAAAGGCCCGGGTGGGCCCTGTGTGCACATGGGGCT	3581
Oy	3457	GCCCCCTGCGCATGACCCCCCAAAAGCTGTGCACTTTACGCGAGATGAAGCAGAGCAGCT	3516
Db	3582	GCCCCCTGCGCATGACCCCCCAAAAGCTGTGCACTTTACGCGAGATGAAGCAGAGCAGCT	3641
Oy	3517	GTCCCAACGGAGGCGAGGCTGGGCCACCGAGAGGCTGTGGGGGTGCGCACGCGCCAGAGGC	3576
Db	3642	GTCCCAACGGAGGCGAGGCTGGGCCACCGAGAGGCTGTGGGGGTGCGCACGCGCCAGAGGC	3701
Oy	3577	GTCCGTGCTGAGAGGACAGCTCTGGGCTCAATTCCGGGCGAAGCATCAACAAAGGCAT	3636

Db 3702 GTCCGTGCTGAGAGGGAACAGCTCTGGGGCTCAGTTCCGGGCGGAAGCATCAACAAAGGAT 3761  
QY 3637 TCCAGCAACAGGGTGGCCCTGGAACAAGCCATCAATACGGGCTCCATCAACCCACGG 3696  
Db 3762 TCCAGCAACAGGGTGGCCCTGGAACAAGCCATCAATACGGGCTCCATCAACCCACGG 3821  
QY 3637 CACGCAAGTCGATCTCTGTATAAGGGCAACATCACAGGAATCATCGGAGAGAACGCC 3756  
Db 3822 CACGCAAGTCGATCTCTGTATAAGGGCAACATCACAGGAATCATCGGAGAGAACGCC 3881  
QY 3757 GAGTCGCTTGAACCGCGCCGGAGAGACAGCTGCGCAAGGGCAAGTCATCTTACGAAG 3816  
Db 3882 GAGTCGCTTGAACCGCGCCGGAGAGACAGCTGCGCAAGGGCAAGTCATCTTACGAAG 3941  
QY 3817 CAAGAAAGGACAGTCTTGTCTATAGGGGTGACATGTCTGTAGCCAGTGTCTCAAGGA 3876  
Db 3942 CAAGAAAGGACAGTCTTGTCTATAGGGGTGACATGTCTGTAGCCAGTGTCTCAAGGA 4001  
QY 3877 GGAAGGAGAGACACTCAAGAACCCGCCATGAGACGGCCGCCCAAGCCGACTATGA 3936  
Db 4002 GGAAGGAGAGACACTCAAGAACCCGCCATGAGACGGCCGCCCAAGCCGACTATGA 4061  
QY 3937 CATGATGAGAGGCGCGGTGGAGAGACCATCTCTCAGCCAGCATGGAAGTCTCATGG 3996  
Db 4062 CATGATGAGAGGCGCGGTGGAGAGACCATCTCTCAGCCAGCATGGAAGTCTCATGG 4121  
QY 4122 CGGTGCTCATCCCGCCGGAGCGACACAGCCCCACCACTCAAGAGACACCAATCCG 4181  
Db 4057 CGGTGCTCATCCCGCCGGAGCGACACAGCCCCACCACTCAAGAGACACCAATCCG 4116  
QY 4182 CGGTGCTCATCCCGCCGGAGCGACACAGCCCCACCACTCAAGAGACACCAATCCG 4241  
Db 4117 TCGGAGAGCGCAAGGATTCCTCTCGGTCTACAGTGAAGGACAGAGAGACTTACCTGG 4176  
QY 4242 TCGGAGAGCGCAAGGATTCCTCTCGGTCTACAGTGAAGGACAGAGAGACTTACCTGG 4301  
Db 4177 TCGGAGAGCGCAAGGATTCCTCTCGGTCTACAGTGAAGGACAGAGAGACTTACCTGG 4236  
QY 4302 CTGACCCGAGGCTCTCAAGACGAGGCCCTGGGCCCTTGAAGCTGAAGCCGGCCATGA 4361  
Db 4237 GGGGCTGTGTGCGCAAGTGAAGAGGCGGGCGGCTTCATTCATGAGATTCGCCGGAGGA 4296  
QY 4362 GGGGCTGTGTGCGCAAGTGAAGAGGCGGGCGGCTTCATTCATGAGATTCGCCGGAGGA 4421  
Db 4297 GCTGCGGCAACAGCCCGAGCTGCCCTGGGCCCGGCGGCTCAAGAGAGGCTCATCAC 4356  
QY 4422 GCTGCGGCAACAGCCCGAGCTGCCCTGGGCCCGGCGGCTCAAGAGAGGCTCATCAC 4481  
Db 4357 GCAAGGCAACCCCGCTCAAGTACGACACCGGCGGCTTCACCACTGACTCCMAAAGACGA 4416  
QY 4482 GCAAGGCAACCCCGCTCAAGTACGACACCGGCGGCTTCACCACTGACTCCMAAAGACGA 4541  
Db 4417 GCTAGGCTCCCTCATTCGGAAGCCCGGCGGCGGCTTCACCACTTCGTCACCCGCTGATGT 4476  
QY 4542 GCTAGGCTCCCTCATTCGGAAGCCCGGCGGCGGCTTCACCACTTCGTCACCCGCTGATGT 4601  
Db 4477 GATGCGCAACGCGCGGCACTGGAAGCTGTGCTACGAGAGAGGCTGAAGAGCGGCGC 4536  
QY 4602 GATGCGCAACGCGCGGCACTGGAAGCTGTGCTACGAGAGAGGCTGAAGAGCGGCGC 4661  
Db 4537 AGGGAACGCGCAAGACTCGGGGGGCTCAATTGCGCGCGGCGCGGCTGATGTGCTGA 4596  
QY 4662 AGGGAACGCGCAAGACTCGGGGGGCTCAATTGCGCGCGGCGCGGCTGATGTGCTGA 4721  
Db 4537 GCTGAGTAAAGCCGCGGCAAGAGCCCTTGAACCTATAGAGACACAGGAGCACTTTGGCGG 4656  
QY 4722 GCTGAGTAAAGCCGCGGCAAGAGCCCTTGAACCTATAGAGACACAGGAGCACTTTGGCGG 4781  
Db 4657 CCACTTCCACAGAGGTTGCGCGGTGACATGCGGAGACCAAGCGCGGCTTCAGAGAGG 4716  
QY 4782 CCACTTCCACAGAGGTTGCGCGGTGACATGCGGAGACCAAGCGCGGCTTCAGAGAGG 4841

QY 4717 CAGCCTTTGCTCAGCAAGGCATTCACAGAACGAAAGCTGAGCTGAGCGCTCTGAGAT 4776  
Db 4842 CAGCCTTTGCTCAGCAAGGCATTCACAGAACGAAAGCTGAGCTGAGCGCTCTGAGAT 4901  
QY 4777 GGCAGAGTCCCGGCAAGACCGGTGCCGAGACCAACCAACCCCATCTGSCCTATGA 4836  
Db 4902 GGCAGAGTCCCGGCAAGACCGGTGCCGAGACCAACCAACCCCATCTGSCCTATGA 4961  
QY 4837 GCACTGTCTTGGGCGTGAAGTGGGGAAGCTGTATCGAGACCAATCCCCGAGCTT 4896  
Db 4962 GCACTGTCTTGGGCGTGAAGTGGGGAAGCTGTATCGAGACCAATCCCCGAGCTT 5021  
QY 4897 GCACTGTCTTGGGCGTGAAGTGGGGAAGCTGTATCGAGACCAATCCCCGAGCTT 4956  
Db 5022 GCACTGTCTTGGGCGTGAAGTGGGGAAGCTGTATCGAGACCAATCCCCGAGCTT 5081  
QY 4957 CCGACACGTGGGCCCAACCCCACTTACCGGACCTGTATCCCACTTACCTATCCGCG 5016  
Db 5082 CCGACACCTGGGCCCAACCCCACTTACCGGACCTGTATCCCACTTACCTATCCGCG 5141  
QY 5017 CTACCCGACACAGCGGCGCTGAGAACCGGACAGACATCATATGACTATCATCTC 5076  
Db 5142 CTACCCGACACAGCGGCGCTGAGAACCGGACAGACATCATATGACTATCATCTC 5201  
QY 5077 GCAAGCATGCAACCAACACGCGCACTGCGCATGGCCCAAGCGAGCTGATATGCTGAAGG 5136  
Db 5202 GCAAGCATGCAACCAACACGCGCACTGCGCATGGCCCAAGCGAGCTGATATGCTGAAGG 5261  
QY 5137 CCTGTGGCCCGCGAGTCTGTGCTGCACTCAATCAGCTGGGGTCCCGAGGATCAT 5196  
Db 5262 CCTGTGGCCCGCGAGTCTGTGCTGCACTCAATCAGCTGGGGTCCCGAGGATCAT 5321  
QY 5197 GCACTGTCTCAAGTGCACACTTGCCTGTGTCTGTGCCCCCGACACAGGACCCGAC 5256  
Db 5322 GCACTGTCTCAAGTGCACACTTGCCTGTGTCTGTGCCCCCGACACAGGACCCGAC 5381  
QY 5257 CACGCGCATGGAACCGGCTTGGCTACCTCCCAAGCGGCGCCGCTTCAGAGGCGGCA 5316  
Db 5382 CACGCGCATGGAACCGGCTTGGCTACCTCCCAAGCGGCGCCGCTTCAGAGGCGGCA 5441  
QY 5317 CAGGAGCTCCCACTCTCCCGAGAGGTCAACACACTTGAACAAACCAACCAACGTC 5376  
Db 5442 CAGGAGCTCCCACTCTCCCGAGAGGTCAACACACTTGAACAAACCAACCAACGTC 5501  
QY 5377 CTGCTCCGAGCGGAGCGAGACCGGATCGAAGCGGAGACCGGATTCGGAGCGGAGAAA 5436  
Db 5502 CTGCTCCGAGCGGAGCGAGACCGGATCGAAGCGGAGACCGGATTCGGAGCGGAGAAA 5561  
QY 5437 GTCCATCTCTCACGTCACACGAGGTGAGACGCAACCATTTGAGACCTGTGTAAGA 5496  
Db 5562 GTCCATCTCTCACGTCACACGAGGTGAGACGCAACCATTTGAGACCTGTGTAAGA 5621  
QY 5497 GCAAGGACCGGCAAGCGGAGGAGGAGGCGGGGTGGGGGCAAGCAAGCGGCGCGC 5556  
Db 5622 GCAAGGACCGGCAAGCGGAGGAGGAGGCGGGGTGGGGGCAAGCAAGCGGCGCGC 5681  
QY 5557 CTCCCACTCCATGCGCAACAGCACTCGCCCATCTCCCTCGAGACCCAGATGCCCTCA 5616  
Db 5682 CTCCCACTCCATGCGCAACAGCACTCGCCCATCTCCCTCGAGACCCAGATGCCCTCA 5741  
QY 5617 GCAAGACCCAGTGTGCTTCAACACAGGATGAAGGATCATCAACCGCTGTGAGGC 5676  
Db 5742 GCAAGACCCAGTGTGCTTCAACACAGGATGAAGGATCATCAACCGCTGTGAGGC 5801  
QY 5677 CAGCAAGCCACAGTCTGAGGTCCACCTCAACCTCTCAACCGTTCGCGCAGCTGCCAC 5736  
Db 5802 CAGCAAGCCACAGTCTGAGGTCCACCTCAACCTCTCAACCGTTCGCGCAGCTGCCAC 5861  
QY 5737 ATTCCCACTGCAACCACTGCGCACTGAGGCGGACCTTCGATGAGGATCTACCTTACCT 5796  
Db 5862 ATTCCCACTGCAACCACTGCGCACTGAGGCGGACCTTCGATGAGGATCTACCTTACCT 5921



5797 CATGAGCCCGCTTGTGCTGCCAAGAGAGCCCCCGGGTCCCGGCGCAAGAGCGCCCG 5856  
5922 CATGAGCCCGCTTGTGCTGCCAAGAGAGCCCCCGGGTCCCGGCGCAAGAGCGCCCG 5981  
5857 AACAGACACCGGCGCATGCTTCTGCGCAAGCGCCCGGCTCCGGGCTGAGAGCCCG 5916  
5982 AACAGACACCGGCGCATGCTTCTGCGCAAGCGCCCGGCTCCGGGCTGAGAGCCCG 6041  
5917 CTCTCTCCCGCAAGAGGCTCGAGAGCCCGGCGCTTATGCTCTGTCTTGGACACG 5976  
6042 CTCTCTCCCGCAAGAGGCTCGAGAGCCCGGCGCTTATGCTCTGTCTTGGACACG 6101  
5977 CACCATCGCCGCAACCTCTGCGAAGAACTTCGACCTGACCAAGCGAGCCCGGAC 6036  
6102 CACCATCGCCGCAACCTCTGCGAAGAACTTCGACCTGACCAAGCGAGCCCGGAC 6161  
6037 GGGCGCACTGCTGCTCGGCTCGGACCCCGCAACCGGAAAGACTCAAAAGTAAACCTTTTC 6096  
6162 GGGCGCACTGCTGCTCGGCTCGGACCCCGCAACCGGAAAGACTCAAAAGTAAACCTTTTC 6221  
6097 CATCCAGAACTGGAACCTCGTTCTCTGGGTTTACACGCGCAGAGCTTACAGCCCGAAG 6156  
6222 CATCCAGAACTGGAACCTCGTTCTCTGGGTTTACACGCGCAGAGCTTACAGCCCGAAG 6281  
6157 GGTGAGCCCGCTCAGCCCTGTGAGCTCACCCAGTTCAGCCAGAGGCGCTCCCA 6216  
6282 GGTGAGCCCGCTCAGCCCTGTGAGCTCACCCAGTTCAGCCAGAGGCGCTCCCA 6341  
6217 GCACTTGAAAGAGCTGACAAAGAGCACTTGGAGGGGAGCTGCGGCCCAAGACCAAG 6276  
6342 GCACTTGAAAGAGCTGACAAAGAGCACTTGGAGGGGAGCTGCGGCCCAAGACCAAG 6401  
6277 CCCCGTGAAGCTTGGGCGGAGGCGCCACCTCCGACCTGCGGCGCTGCTGAG 6336  
6402 CCCCGTGAAGCTTGGGCGGAGGCGCCACCTCCGACCTGCGGCGCTGCTGAG 6461  
6337 CCAGCCCTGTCCAGCCCGCTGCTCCAGACCGCCCGGAGGTCAAGAGTACCGAGCGGT 6396  
6462 CCAGCCCTGTCCAGCCCGCTGCTCCAGACCGCCCGGAGGTCAAGAGTACCGAGCGGT 6521  
6397 GGTCACTCTGCGCCAGACATATGTAAGGTCAATCACAGAGTACACCCGACCAACC 6456  
6522 GGTCACTCTGCGCCAGACATATGTAAGGTCAATCACAGAGTACACCCGACCAACC 6581  
6457 ACAGAGCTCAGAGCAACCCCTGCGCGCCCGCTTACTCTTCCCTGGGAGCAGCTGCC 6516  
6582 ACAGAGCTCAGAGCAACCCCTGCGCGCCCGCTTACTCTTCCCTGGGAGCAGCTGCC 6641  
6517 CGTCTGGAACCTTCGCGCGCCCAACCAAGTACCTTACCTTCGCGCCCGGACCAATGCTC 6576  
6642 CGTCTGGAACCTTCGCGCGCCCAACCAAGTACCTTACCTTCGCGCCCGGACCAATGCTC 6701  
6577 CCCGCGCCCTGCTCTCCCGCAAGCGAAGGCGCAAGAGTCTCCAGAGCCAAACAAGC 6636  
6702 CCCGCGCCCTGCTCTCCCGCAAGCGAAGGCGCAAGAGTCTCCAGAGCCAAACAAGC 6761  
6637 GTCGCTTGGGAGTGTGTAAGGAGTGAATTGAACCTGTGCCCAACCGGAGGCAATGAC 6696  
6762 GTCGCTTGGGAGTGTGTAAGGAGTGAATTGAACCTGTGTCCCAACCGGAGGCAATGAC 6821  
6697 GAGAGCCAGGCACTCCGAGAGTGTGTATCCGCTGCTGTACCGGAGTGGGAAACAGC 6756  
6822 GAGAGCCAGGCACTCCGAGAGTGTGTATCCGCTGCTGTACCGGAGTGGGAAACAGC 6881  
6757 GAGAGCCAGAGATGGGCTTCAAGTCTCCAGGCAACCAAGCAGCCGCTTCTT 6816  
6882 GAGAGCCAGAGATGGGCTTCAAGTCTCCAGGCAACCAAGCAGCCGCTTCTT 6941  
6817 CAGCAAGCTGACCGAGAGCACTCCGCAATGCTCAAGTCAAGAGCAAGATCAACA 6876  
6942 CAGCAAGCTGACCGAGAGCACTCCGCAATGCTCAAGTCAAGAGCAAGATCAACA 7001  
6877 GAAGCTGAACACCCAAACCGGAATGAGCCTGAATTAATATCAGCAGCCTGGAGCGA 6936

7002 GAAGCTGAACACCCAAACCGGAATGAGCCTGAATTAATATCAGCAGCCTGGAGCGA 7061  
6937 GATCTTCAATATGCGCGCATCAGCGAAACAGGCTTATGACTTATAGAACAGCGGT 6996  
7062 GATCTTCAATATGCGCGCATCAGCGAAACAGGCTTATGACTTATAGAACAGCGGT 7121  
6997 GCAGGAACATGCGCAGCAACCAATGAGGCTGAGAGGCATTAATTAGAAAGCACTATGG 7056  
7122 GCAGGAACATGCGCAGCAACCAATGAGGCTGAGAGGCATTAATTAGAAAGCACTATGG 7181  
7057 TAAATATGACCAAGTGGAAAGATCCCCCGGCTCAGCGCAATGCTTTTAAACCTTGA 7116  
7182 TAAATATGACCAAGTGGAAAGATCCCCCGGCTCAGCGCAATGCTTTTAAACCTTGA 7241  
7117 TGCAGTGCAGCCTGCGCGCTGCTATGCTTAACCGCTGCTGACGAGAGTACCA 7176  
7242 TGCAGTGCAGCCTGCGCGCTGCTATGCTTAACCGCTGCTGACGAGAGTACCA 7301  
7177 CACACTCACTCGCAAGTGGCGGCGGAAAGCCAAAGTCTTGGCAGACCCAGCAGCG 7236  
7302 CACACTCACTCGCAAGTGGCGGCGGAAAGCCAAAGTCTTGGCAGACCCAGCAGCG 7361  
7237 AAAAGCAAGTCCCGGCGCTGCGCATCTGGGGAACCGGCCACCTCTGTCTCTC 7296  
7362 AAAAGCAAGTCCCGGCGCTGCGCATCTGGGGAACCGGCCACCTCTGTCTCTC 7421  
7297 AGTGAACCTCGAGGAGAGCTGCAACCGCGGAGCGGCTCACCAACCGGCTGGGAGGA 7356  
7422 AGTGAACCTCGAGGAGAGCTGCAACCGCGGAGCGGCTCACCAACCGGCTGGGAGGA 7481  
7357 CAGGCCCTGTCCGAGGTTCCAGCGCAATCCCTTAACACCCCTGATCATCGCTGCA 7416  
7482 CAGGCCCTGTCCGAGGTTCCAGCGCAATCCCTTAACACCCCTGATCATCGGCTGCA 7541  
7417 GCGGAGTGTATGCTTCCCAACCCCGGCGCTCCCGGCGGAGCGGCGCTTCCG 7476  
7542 GCGGAGTGTATGCTTCCCAACCCCGGCGCTCCCGGCGGAGCGGCGCTTCCG 7601  
7477 TGGCCCCCAACGAGCTGGGAGAGGAGGCCAAGGCCACTGTGTCAGATACAGAGC 7536  
7602 TGGCCCCCAACGAGCTGGGAGAGGAGGCCAAGGCCACTGTGTCAGATACAGAGC 7661  
7537 ACTCTCCAGCAGCGAGTGACTCAGAACAGGCGGAGGAGGAGCGGTGEBUSPERLUG 7596  
7662 ACTCTCCAGCAGCGAGTGACTCAGAACAGGCGGAGGAGGAGCGGTGEBUSPERLUG 7711  
7597 TCAAGTCCAGCGAGCACAAGAAAGGCTTCAAGAGAGCGGCGCTGCCACTCCCC 7656  
7712 TCAAGTCCAGCGAGCACAAGAAAGGCTTCAAGAGAGCGGCGCTGCCACTCCCC 7771  
7657 AACCAAGAAAGAGCCCGTGAATCCGCGTGGCGCTCCATCAATCTGTCGACAGCGG 7716  
7772 AACCAAGAAAGAGCCCGTGAATCCGCGTGGCGCTCCATCAATCTGTCGACAGCGG 7831  
7717 GCATCTTGCCTGTCTAAAGCCTTAAGACTCCGCGCCGAGCTGAGCCTGTGACA 7776  
7832 GCATCTTGCCTGTCTAAAGCCTTAAGACTCCGCGCCGAGCTGAGCCTGTGACA 7891  
7777 CCTTACTCAGGAGATGTTTACTGTGCTCGGAAAGGAGGAGGAGCGCGGAGG 7836  
7892 CCTTACTCAGGAGATGTTTACTGTGCTCGGAAAGGAGGAGGAGCGCGGAGG 7951  
7837 GGAAGCGAGGCGTGGGCGAGCAACACAGGCGGCGAGGCGGCAAGGACCAAAAGCAG 7896  
7952 GGAAGCGAGGCGTGGGCGAGCAACACAGGCGGCGAGGCGGCAAGGACCAAAAGCAG 8011  
7897 GATGACCAAGCACTCAAGCGCACTGCGCTCCCGAATGCAATTTGGAAACAAAGTCTTAA 7956  
8012 GATGACCAAGCACTCAAGCGCACTGCGCTCCCGAATGCAATTTGGAAACAAAGTCTTAA 8071  
7957 CTGAGCTGCAAGCCCGCGCTCTCCCTGCGCTCCATCCGCTTAAGGCTCTGACAG 8016



Db 8072 CTGAGCTGCAGCCCCCGGCGCTCCCTCCGCTCCCATCCCGCTTAAAGCGCTGTGAACG 8131  
QY 8017 ATGAGCGAGGCGCTGTCCAGGCCCCAGTGCCTGTCCGGTCCCAACAAGTCCCA 8076  
Db 8132 ATGAGCGAGGCGCTGTCCAGGCCCCAGTGCCTGTCCGGTCCCAACAAGTCCCA 8191  
QY 8077 GCCAAGGATTTGCTGGAACCAAGTCAGGCGAGGTGGCGGACAAAAGGCGAGTGC 8136  
Db 8192 GCCAAGGATTTGCTGGAACCAAGTCAGGCGAGGTGGCGGACAAAAGGCGAGTGC 8251  
QY 8137 GCGTGGGGGGAACGAGATGCTCCGAGAGCTGAGCTGTCTTTTTCACACATCGTTCGAG 8196  
Db 8252 GCTGGGGGGAACGAGATGCTCCGAGAGCTGAGCTGTCTTTTTCACACATCGTTCGAG 8311  
QY 8197 CGGTGGGAAGGAAGGAGAGTGAATGATGTGTGTTTACAGGGTATATTTTGTATAC 8256  
Db 8312 CGGTGGGAAGGAAGGAGAGTGAATGATGTGTGTTTACAGGGTATATTTTGTATAC 8371  
QY 8257 CTTCAATGAATTAATTCAGATGTTTACGCAAGAGAGCTTACCCAGTATTACTGTGC 8316  
Db 8372 CTTCAATGAATTAATTCAGATGTTTACGCAAGAGAGCTTACCCAGTATTACTGTGC 8431  
QY 8317 TGTGCTTTTGAATCTGTGCTTACCGTTCAAGAGCGGTGCGAGCGAGAGTGCAGACC 8376  
Db 8432 TGTGCTTTTGAATCTGTGCTTACCGTTCAAGAGCGGTGCGAGCGAGAGTGCAGACC 8491  
QY 8377 CATCATCGCAGAGCAAGAGGCGGAGCTGCTGTCACGCGCGCTGTGTCTCCCTC 8436  
Db 8492 CATCATCGCAGAGCAAGAGGCGGAGCTGCTGTCACGCGCGCTGTGTCTCCCTC 8551  
QY 8437 CCTCCCTTCTTGGGAGAGTGAATTCGATGCGTATCTGTGGCGGCAATTTGGCGAGG 8496  
Db 8552 CCTCCCTTCTTGGGAGAGTGAATTCGATGCGTATCTGTGGCGGCAATTTGGCGAGG 8611  
QY 8457 TGTGCTTTTGAATCTGTGCTTACCGTTCAAGAGCGGTGCGAGCGAGAGTGCAGACC 8556  
Db 8612 TGTGCTTTTGAATCTGTGCTTACCGTTCAAGAGCGGTGCGAGCGAGAGTGCAGACC 8671  
QY 8557 AAAAAAAAAAAAAA 8571  
Db 8672 AAAAAAAAAAAAAA 8686

RESULT 6  
US-10-723-860-1739  
; Sequence 1739, Application US/10723860  
; Publication No. US20040253606A1  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Gineburg, Wendy M.  
; APPLICANT: Zlotnick, Albert  
; TITLE OF INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions &  
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators  
; FILE REFERENCE: 05882, 0193, NPO501  
; CURRENT APPLICATION NUMBER: US/10/723, 860  
; PRIOR FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: 60/429, 739  
; NUMBER OF SEQ ID NOS: 8393  
; SOFTWARE: Patent version 3.2  
; SEQ ID NO 1739  
; LENGTH: 8686  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-723-860-1739

Query Match 97.2%; Score 8324.6; DB 8; Length 8686;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 8472; Conservative 1; Mismatches 34; Indels 88; Gaps 4;

QY 1 CATGTGGGCTTCACACAGCTTTGTGGCAGACGTTGAGGCGCCACTGAGCCCGCTTACCC 60  
Db 156 CATGTGGGCTTCACACAGCTTTGTGGCAGACGTTGAGGCGCCACTGAGCCCGCTTACCC 215

QY 61 GCCCAGAGCCTTTCTTACCCAGTGCAGATGCTCCGAGCGACACGAGCTGCGGCTCTT 120  
Db 216 GCCCAGAGCCTTTCTTACCCAGTGCAGATGCTCCGAGCGACACGAGCTGCGGCTCTT 275  
QY 121 GAGATTCACAGACCATCCCGGAGCTATGCTCTCCACCTGTGCGCGGCTCTCATCATCA 180  
Db 276 GAGATTCACAGACCATCCCGGAGCTATGCTCTCCACCTGTGCGCGGCTCTCATCATCA 335  
QY 181 GCCCAGAGCGAGGAGCCTCTCTGCTGTGAGTTCAGCTCCGAGATGAACGATGCCA 240  
Db 336 GCCCAGAGCGAGGAGCCTCTCTGCTGTGAGTTCAGCTCCGAGATGAACGATGCCA 395  
QY 241 GAGCTTCACCTGCGGCGAGAGTCCACTCATPACTGCTCCGAGCTGGGAGATCAGAGAT 300  
Db 366 GAGCTTCACCTGCGGCGAGAGTCCACTCATPACTGCTCCGAGCTGGGAGATCAGAGAT 455  
QY 301 GAGTTCATTTGAAGCAAGCCCTCCGCTGAGTGTGCTGACCCCTGCGGAGC 360  
Db 456 GAGTTCATTTGAAGCAAGCCCTCCGCTGAGTGTGCTGACCCCTGCGGAGC 515  
QY 361 GTACACCTGCTGCGCAGCGGCTGAGCTGCGGAGTCTGAAGACTCAACAAGACCTGAG 420  
Db 516 GTACACCTGCTGCGCAGCGGCTGAGCTGCGGAGTCTGAAGACTCAACAAGACCTGAG 575  
QY 421 CTTGACGCGGCAAGCTGGAACCGGTGTCTCCCTCCAGCCCTCCGCACTGACCTGAGCT 480  
Db 576 CTTGACGCGGCAAGCTGGAACCGGTGTCTCCCTCCAGCCCTCCGCACTGACCTGAGCT 635  
QY 481 GAGCTGTGTGCGGCTGAGCTGTCTCCAGAGAGAGTGTATCCAGAACATGAGACCGGTGGA 540  
Db 636 GAGCTGTGTGCGGCTGAGCTGTCTCCAGAGAGAGTGTATCCAGAACATGAGACCGGTGGA 695  
QY 541 CCGAGAGATTCACATGTGTGAGAGCAGACAGATCTTAAGCTGAAGAGAGAGACAGACT 600  
Db 696 CCGAGAGATTCACATGTGTGAGAGCAGACAGATCTTAAGCTGAAGAGAGAGACAGACT 755  
QY 601 GAGAGAGAGGCTGCGCAAGCCGCGGAGCTGAGAGCCGCTGTGACCCGCGCATGGA 660  
Db 756 GAGAGAGAGGCTGCGCAAGCCGCGGAGCTGAGAGCCGCTGTGACCCGCGCATGGA 815  
QY 661 GTCCAGACCCGCGAGCTGTGTGAGATATCTTACAGAGAGAACCGGAGAGGCTGAAGC 720  
Db 816 GTCCAGACCCGCGAGCTGTGTGAGATATCTTACAGAGAGAACCGGAGAGGCTGAAGC 875  
QY 721 TGACATGAGATTCGGAAGGCTGAGGCGCCAGGTGAGCTGCGCTGTCAACACAGC 780  
Db 876 TGACATGAGATTCGGAAGGCTGAGGCGCCAGGTGAGCTGCGCTGTCAACACAGC 935  
QY 781 CTCCGACACCCGCGAGATTCATGAGAACATGAATAAACAGGCGATGCGAGAGAGCT 840  
Db 936 CTCCGACACCCGCGAGATTCATGAGAACATGAATAAACAGGCGATGCGAGAGAGCT 995  
QY 841 AATCTTGTATCTTCAAGAGAGAGATCAGCTGCGAGAACATGAGAGAGAGAGTTCGCA 900  
Db 996 AATCTTGTATCTTCAAGAGAGAGATCAGCTGCGAGAACATGAGAGAGAGAGTTCGCA 1055  
QY 901 GCGCTATGACAGCTCATGAGAGGCTTGGAAAAAAGGTGAGCGCATGCAAAACACCC 960  
Db 1056 GCGCTATGACAGCTCATGAGAGGCTTGGAAAAAAGGTGAGCGCATGCAAAACACCC 1115  
QY 961 GCGCGCGCGGCGCAAGAGACCAAGGTGCGAGTATCTACAAAAAGCATTCCTGAGAT 1020  
Db 1116 GCGCGCGCGGCGCAAGAGACCAAGGTTCGAGATCTACAGAGAGAGGTTCCGAGAT 1175  
QY 1021 CCGAGAGAGCGGAGCTGACAGAGCGATGCAAGGCTGAGCGGCGAGCGGAGCGTGG 1080  
Db 1176 CCGAGAGAGCGGAGCTGACAGAGCGATGCAAGGCTGAGCGGCGAGCGGAGCGTGG 1232  
QY 1081 GCTGTTCATGTGCGCGCGCGGCGAGACAGAGGTGTCAAGAGATCATGATGCGCTGC 1140  
Db 1233 GCTGTTCATGTGCGCGCGCGGCGAGACAGAGGTGTCAAGAGATCATGATGCGCTGC 1292

1141 AGAGCAGAGAACTTGAGAGAGAGATGCGCCAGCTGCGCCGTATCCGCCATGCTGTAA 1200  
1293 AGAGCAGAGAACTTGAGAGAGAGATGCGCCAGCTGCGCCGTATCCGCCATGCTGTAA 1352  
1201 CGAGCTTGACACAGACGCGCATCAAGTTTCATCAATGAACGCGCTTATGCGCCGAT 1260  
1353 CGAGCTTGACACAGACGCGCATCAAGTTTCATCAATGAACGCGCTTATGCGCCGAT 1412  
1261 GAAGGTGTCAAAAGACCGCAGGCTCATGAAATGTGAGTGAAGAGAGAGAGAGAGAGAG 1320  
1413 GAAGGTGTCAAAAGACCGCAGGCTCATGAAATGTGAGTGAAGAGAGAGAGAGAGAGAG 1472  
1321 CGGAGAAAGTTTCATGACAGATCCCAAGAACTTTGCGCTGATGCAATCTCTGAGAG 1380  
1473 CGGAGAAAGTTTCATGACAGATCCCAAGAACTTTGCGCTGATGCAATCTCTGAGAG 1532  
1381 GAAGACATGCGCTGAGTGCCTCTTATTAATACTGACTAAGAGATGAGAACTATTA 1440  
1533 GAAGACATGCGCTGAGTGCCTCTTATTAATACTGACTAAGAGATGAGAACTATTA 1592  
1441 GAGCTGTGAGAGCGAGGCTATGCGCGCGCGCAAGAGCGAGAGCAACAAGAGAGCA 1500  
1593 GAGCTGTGAGAGCGAGGCTATGCGCGCGCGCAAGAGCGAGAGCAACAAGAGAGCA 1652  
1501 GAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
1653 GAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1712  
1561 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
1713 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1772  
1621 CGACAAAGAGAACTTCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
1773 CGACAAAGAGAACTTCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1832  
1681 GAAGAGAGCTGTGCGCTCCAAAGAGCGCAAACTGCAACAGCGAGAGAGAGAGAGAG 1740  
1833 GAAGAGAGCTGTGCGCTCCAAAGAGCGCAAACTGCAACAGCGAGAGAGAGAGAGAG 1892  
1741 CGGCATCACCCGCTCAATGAGCTTAATGAGGCAACAGCGAGAGAGAGAGAGAGAG 1800  
1893 CGGCATCACCCGCTCAATGAGGCTTAATGAGGCAACAGCGAGAGAGAGAGAGAGAG 1952  
1801 GAGCGCGGAGCTGCGCTTCAATGAGGCTTAATGAGGCTTCCGCTGAGAGAGAGAGAG 1860  
1953 GAGCGCGGAGCTGCGCTTCAATGAGGCTTAATGAGGCTTCCGCTGAGAGAGAGAGAG 2012  
1861 GGAAGACAGCAAGAAAGTCTCTGAGAAACAGCGCGCAACTGAGTGGCGCATCGCCGAT 1920  
2013 GGAAGACAGCAAGAAAGTCTCTGAGAAACAGCGCGCAACTGAGTGGCGCATCGCCGAT 2072  
1921 GGTGGGCTTCAAGACTGTGTGCGAGTGTAAAGAACTTTCACTTCAACTACAAAGAGAGCA 1980  
2073 GGTGGGCTTCAAGACTGTGTGCGAGTGTAAAGAACTTTCACTTCAACTACAAAGAGAGCA 2132  
1981 GAACCTCGATGAGATTTTGACAGACACAAGCTGAAAGATGAGAGAGAGAGAGAGAGAG 2040  
2133 GAACCTCGATGAGATTTTGACAGACACAAGCTGAAAGATGAGAGAGAGAGAGAGAGAG 2192  
2041 GAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
2193 GAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2252  
2101 GAGAGATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160  
2253 GAGAGATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2312  
2161 GGCTGAGAGCTTATGATGCTCTGAGAGATGAGAGTGGCCAGAGAGAGAGATGAGTGGCCAGC 2220  
2313 GGCTGAGAGCTTATGATGCTCTGAGAGATGAGAGTGGCCAGAGAGAGAGATGAGTGGCCAGC 2320  
2221 CACTGTCAACAACAGCTCAGACACAGAGAGATCCCTCTCTCACTGAGAGCGCCAA 2280

2322 CACTGTCAACAACAGCTCAGACACAGAGAGATCCCTCTCTCACTGAGAGCGCCAA 2381  
2281 GAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
2382 GAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2441  
2341 AGGCGCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCA 2400  
2442 AGGCGCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCA 2501  
2401 TGAAGCCACCGAGAGCCCTTAAGCCCTCAACAGACATCCCAATGCGCTTGTACATCTTCC 2460  
2502 TGAAGCCACCTTAAGCCCTTAAGCCCTCAACAGACATCCCAATGCGCTTGTACATCTTCC 2561  
2461 TGTGTGCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
2562 TGTGTGCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2621  
2521 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
2622 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2681  
2581 GCCCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
2682 GCCCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2741  
2641 GAGAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
2742 GAGAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2801  
2701 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760  
2802 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2861  
2761 CTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
2862 CTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2921  
2821 AAGGCGCAAGCTTCTCAACCCCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
2922 AAGGCGCAAGCTTCTCAACCCCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2981  
2881 ACTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
2982 ACTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3041  
2941 AGTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000  
3042 AGTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3101  
3001 GCCACGCAAAACCTGACAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
3102 GCCACGCAAAACCTGACAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3161  
3061 GGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3100  
3162 GGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3221  
3101 ---GCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3156  
3222 CCAAGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3281  
3157 CCGGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3216  
3282 CCGGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3341  
3217 CTGAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3276  
3342 CTGAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3401  
3277 CCGGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3336

Db 3402 CGGCGCGTCTGCGCGCGCCACCCACCATCTCCAAACCGGCTCCCTCATCTCTGTC 3461  
Qy 3337 CAAGCACCACGAGCTCTTGAGAGGCAATAAGTGCATCTCCAAAGATGTGGTCCA 3396  
Db 3462 CAAGCACCACGAGCTCTTGAGAGGCAATAAGTGCATCTCCAAAGATGTGGTCCA 3521  
Qy 3397 GGTCCAGCTCCCGTACTCAGAGCATGCGAAGGCCCGGCGGCGCTGTCAACATGGGGCT 3456  
Db 3522 GCTCCAGCTCCCGTACTCAGAGCATGCGAAGGCCCGGCGGCGCTGTCAACATGGGGCT 3581  
Qy 3457 GCCCTGCGCATGAGCCCCAAAAGCTGGCACCTTTAGCGGAGTGAAGCAGAGCAGCT 3516  
Db 3582 GCCCTGCGCATGAGCCCCAAAAGCTGGCACCTTTAGCGGAGTGAAGCAGAGCAGCT 3641  
Qy 3517 GTCCCAACGGGGCAGGCTGGGCAACCGAGAGCTTGGGGGTGCCACAGCCCAAGAGGC 3576  
Db 3642 GTCCCAACGGGGCAGGCTGGGCAACCGAGAGCTTGGGGGTGCCACAGCCCAAGAGGC 3701  
Qy 3577 GTCCGTGCTGAGAGGGAAGGCTGTGGGCTCAGTTCCGGGCGGAAGCATCAACAAAGCAT 3636  
Db 3702 GTCCGTGCTGAGAGGGAAGGCTGTGGGCTCAGTTCCGGGCGGAAGCATCAACAAAGCAT 3761  
Qy 3637 TCCCAAGCAACGGGGTCCCTCGGACAGGCGCATCAATACGGGGTCCATCACCGG 3696  
Db 3782 TCCCAAGCAACGGGGTCCCTCGGACAGGCGCATCAATACGGGGTCCATCACCGG 3821  
Qy 3697 CACGCGACGTAGCGTCTGTGACAAAGGCAACATCAACAGATCATCGGAGAGACGCC 3756  
Db 3822 CACGCGACGTAGCGTCTGTGACAAAGGCAACATCAACAGATCATCGGAGAGACGCC 3881  
Qy 3757 GAGTGGCTTTGGAACCGCGCGCGGGAAGACAGCTGCGCCAAAGGCGCATCTTACGAAGG 3816  
Db 3882 GAGTGGCTTTGGAACCGCGCGCGGGAAGACAGCTGCGCCAAAGGCGCATCTTACGAAGG 3941  
Qy 3817 CAAGAAAGGAGCCAGTCTTGTCTTATGAGGGTGGCAATGTGTGACCGAGTGTCCAAAGG 3876  
Db 3942 CAAGAAAGGAGCCAGTCTTGTCTTATGAGGGTGGCAATGTGTGACCGAGTGTCCAAAGG 4001  
Qy 3877 GAGCGGAGAAAGCAGCTCAGACCCGCCCATGAGACGGCGCGCCCAAGCGACCTATGA 3936  
Db 4002 GAGCGGAGAAAGCAGCTCAGACCCGCCCATGAGACGGCGCGCCCAAGCGACCTATGA 4061  
Qy 3937 CATGATGAGAGGGCGCGTGGGCAAGAGCCATCTCTCAAGCAGCATCGAAGGTCTATGGG 3996  
Db 4062 CATGATGAGAGGGCGCGTGGGCAAGAGCCATCTCTCAAGCAGCATCGAAGGTCTATGGG 4121  
Qy 3997 CGGTCATCCCGCGGAGCGACAGGCGCCGACCATCAAGAGCGAGCACATCGG 4056  
Db 4122 CGGTCATCCCGCGGAGCGACAGGCGCCGACCATCAAGAGCGAGCACATCGG 4181  
Qy 4057 CGGTCATCAACAAGGATCCCTCGTCTTACGTGAGGAGCAGAGAGCTACTGGG 4116  
Db 4182 CGGTCATCAACAAGGATCCCTCGTCTTACGTGAGGAGCAGAGAGCTACTGGG 4241  
Qy 4117 TCGGAGAGCCAAAGTCTTAAAGCGGAGGAGCAGCTTCGCCCCACCGGCTTACGGGA 4176  
Db 4242 TCGGAGAGCCAAAGTCTTAAAGCGGAGGAGCAGCTTCGCCCCACCGGCTTACGGGA 4301  
Qy 4177 CTTGACCGAGGCTTCAAGAGCGCAGGCGCTTGGGCCCTTGAAGCTGAAGCGGCGCATGA 4236  
Db 4302 CTTGACCGAGGCTTCAAGAGCGCAGGCGCTTGGGCCCTTGAAGCTGAAGCGGCGCATGA 4361  
Qy 4237 GGGCGTGTGGCCAAGTGAAGAGGCGGCGCTCTCATTCATGAGATCCCGCGAAGGA 4296  
Db 4362 GGGCGTGTGGCCAAGTGAAGAGGCGGCGCTCTCATTCATGAGATCCCGCGAAGGA 4421  
Qy 4297 GCTGGGGCAACGCGCCGAGCTGCCCTTGGCCCCCGGGCGGCTCAAGAGGGCTTCATCAC 4356  
Db 4422 GCTGGGGCAACGCGCCGAGCTGCCCTTGGCCCCCGGGCGGCTCAAGAGGGCTTCATCAC 4481  
Qy 4357 GCAGGGCAACCGGCTCAAGTGAAGACCGGGCGGCTCAACCTGGCTCAAAAAGACAGA 4416  
Db 4482 GCAGGGCAACCGGCTCAAGTGAAGACCGGGCGGCTCAACCTGGCTCAAAAAGACAGA 4541

Qy 4417 GGTACGCTTCCTCATTCGAGAGCGCCCGGCGGACGTTCCACCCGTGACCCGCTGATGT 4476  
Db 4542 GGTACGCTTCCTCATTCGAGAGCGCCCGGCGGACGTTCCACCCGTGACCCGCTGATGT 4601  
Qy 4477 GATGGCGCAAGCCCGGGGACGTGGAACGTGCGCTGCTCAAGAGAGAGCTGAAGAGCGGGC 4536  
Db 4602 GATGGCGCAAGCCCGGGGACGTGGAACGTGCGCTGCTCAAGAGAGAGCTGAAGAGCGGGC 4661  
Qy 4537 AGGACCGCCAGAGCTCGGGGGGCTCATTTGCGGCGGAGCGCCCGGCTCAATGTGCTGA 4596  
Db 4682 AGGACCGCCAGAGCTCGGGGGGCTCATTTGCGGCGGAGCGCCCGGCTCAATGTGCTGA 4721  
Qy 4597 GCTGGTAAAGCTCGCGGACAGACCCCTTGAACCTATGAGACCAACGGGACACCTTTGCCGG 4656  
Db 4722 GCTGGTAAAGCTCGCGGACAGACCCCTTGAACCTATGAGACCAACGGGACACCTTTGCCGG 4781  
Qy 4657 CCACTTCCCAAGAGTTGGCCCGTGAACATGCGGGAAGCCACGCGGCGCTGAGAGAGG 4716  
Db 4782 CCACTTCCCAAGAGTTGGCCCGTGAACATGCGGGAAGCCACGCGGCGCTGAGAGAGG 4841  
Qy 4717 CAGCTTTGCTCAGCAAGGATCCAGAACGGAAGCTGACGTGACGCTCGTGAAT 4776  
Db 4842 CAGCTTTGCTCAGCAAGGATCCAGAACGGAAGCTGACGTGACGCTCGTGAAT 4901  
Qy 4777 CGCAGTTCCTCGCACAGACCGTGCAGAGACCAACCAACCCCATCTGCGCTATGA 4836  
Db 4902 CGCAGTTCCTCGCACAGACCGTGCAGAGACCAACCAACCCCATCTGCGCTATGA 4961  
Qy 4837 GCACTGTGTTGGGGCGTGAAGTGGGGTGAACCTGATATGAGAGCAATCCCGTGGGCTT 4896  
Db 4962 GCACTGTGTTGGGGCGTGAAGTGGGGTGAACCTGATATGAGAGCAATCCCGTGGGCTT 5021  
Qy 4897 GCACCCCACTTCATACCCCGGAGATCCCTGTGAGCAGACCGGTGCTTACTACTGCC 4956  
Db 5022 GCACCCCACTTCATACCCCGGAGATCCCTGTGAGCAGACCGGTGCTTACTACTGCC 5081  
Qy 4957 CGACACCTGGCCCCCAACCCCAACCTTACCGGACCTGTATCCCACTTACTTCGCGG 5016  
Db 5082 CGACACCTGGCCCCCAACCCCAACCTTACCGGACCTGTATCCCACTTACTTCGCGG 5141  
Qy 5017 CTACCCCGACACGGGGGGCTGGAAGACGGGACGACATCAATCAATGACTTACTACTTC 5076  
Db 5142 CTACCCCGACACGGGGGGCTGGAAGACGGGACGACATCAATCAATGACTTACTACTTC 5201  
Qy 5077 GCAGCAGATGACCAACAACAGGCAACCGCATGGCCAGAGAGCTGATATGAGGGG 5136  
Db 5202 GCAGCAGATGACCAACAACAGGCAACCGCATGGCCAGAGAGCTGATATGAGGGG 5261  
Qy 5137 CTTCTGCGCCGCGAGTCTTGGCTGGCACTCAACTAGCTGCGGATCCCGAGGATCAT 5196  
Db 5262 CTTCTGCGCCGCGAGTCTTGGCTGGCACTCAACTAGCTGCGGATCCCGAGGATCAT 5321  
Qy 5197 GCACTGTGCCAAGTGGCAACCTGCGTGTGCTGTGCGCCCGACACAGGACCCGACG 5256  
Db 5322 GCACTGTGCCAAGTGGCAACCTGCGTGTGCTGTGCGCCCGACACAGGACCCGACG 5381  
Qy 5257 CACCGCATGAGACGCGCTTGGCTTACTTCCACCGGCGCCAGAGCTTCAAGAGCGGCA 5316  
Db 5382 CACCGCATGAGACGCGCTTGGCTTACTTCCACCGGCGCCAGAGCTTCAAGAGCGGCA 5441  
Qy 5317 CAGCAGCTTCCCACTTCTCCAGAGAGTCAACAACCTTGAACAAAACCAACCAACGTC 5376  
Db 5442 CAGCAGCTTCCCACTTCTCCAGAGAGTCAACAACCTTGAACAAAACCAACCAACGTC 5501  
Qy 5377 CTTCGTCGAAGGGAGGAGACCGGATCGAGAGCGGGAACCGGATTCGGAGCGGGAATA 5436  
Db 5502 CTTCGTCGAAGGGAGGAGACCGGATCGAGAGCGGGAACCGGATTCGGAGCGGGAATA 5561  
Qy 5437 GTCCATCTCTCACTTCAACGACGATGGAACAGCAACCACTTGAAGACCTGTATCAGA 5496  
Db 5562 GTCCATCTCTCACTTCAACGACGATGGAACAGCAACCACTTGAAGACCTGTATCAGA 5621



Db 7772 AACCAAGAGAGCCCTGAGTCCGCTCCATCATCTGCTCCAGAGCCG 7831  
QY 7717 GCATCTTGGCTGTCTAAAGCCTTAAGTCTCCGCGCCGCGGCTGGCTGTGAGA 7776  
Db 7832 GCATCTTGGCTGTCTAAAGCCTTAAGTCTCCGCGCCGCGGCTGGCTGTGAGA 7891  
QY 7777 CTTTACGAGGAGATGTTTACCTGTGCTCGGAAAGGAGGAGGCGGAGGCG 7836  
Db 7892 CTTTACGAGGAGATGTTTACCTGTGCTCGGAAAGGAGGAGGCGGAGGCG 7951  
QY 7837 GCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7896  
Db 7952 GCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8011  
QY 7897 GATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7956  
Db 8012 GATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8071  
QY 7957 CTGAGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8016  
Db 8072 CTGAGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8131  
QY 8017 ATGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8076  
Db 8132 ATGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8191  
QY 8077 GCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8136  
Db 8192 GCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8251  
QY 8137 GCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8196  
Db 8252 GCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8311  
QY 8197 GCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8256  
Db 8312 GCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8371  
QY 8257 CTTCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8316  
Db 8372 CTTCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8431  
QY 8317 TGTGCTTTGATCTCTGCTTACCGTTCAAGAGGCGGTGACAGGCGGAGGAG 8376  
Db 8432 TGTGCTTTGATCTCTGCTTACCGTTCAAGAGGCGGTGACAGGCGGAGGAG 8491  
QY 8377 CATGACTCGCAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8436  
Db 8492 CATGACTCGCAGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8551  
QY 8437 CTTCCCTTCCTTGGGAGAGATGATTCAGATGATTCAGATGATTCAGATGATTCAG 8496  
Db 8552 CTTCCCTTCCTTGGGAGAGATGATTCAGATGATTCAGATGATTCAGATGATTCAG 8611  
QY 8497 TGTGCTTTGATCTCTGCTTACCGTTCAAGAGGCGGTGACAGGCGGAGGAG 8556  
Db 8612 TGTGCTTTGATCTCTGCTTACCGTTCAAGAGGCGGTGACAGGCGGAGGAG 8671  
QY 8557 AAAAAAAAAAAAAA 8571  
Db 8672 AAAAAAAAAAAAAA 8696

## RESULT 7

US-10-723-860-6114/c  
; Sequence 6114, Application US/10723860  
; Publication No. US20040253606A1  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Gineburg, Wendy M.  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions &  
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators

FILE REFERENCE: 05882.0193.NPUS01  
; CURRENT APPLICATION NUMBER: US/10/723.860  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: 60/429,739  
; PRIOR FILING DATE: 2002-11-26  
; NUMBER OF SEQ ID NOS: 8393  
; SOFTWARE: Patent version 3.2  
; SEQ ID NO 6114  
; LENGTH: 9079  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-723-860-6114

Query Match 96.0%; Score 8222.2; DB 8; Length 9079;

Best Local Similarity 98.1%; Pred. No. 0;  
Matches 8435; Conservative 1; Mismatches 68; Indels 93; Gaps 8;

QY 1 CATGTGCGGCTTCACACAGCTTTGTGGACACAGCTGAGAGGCGCACTGAGCCCGCTTACC 60  
Db 8553 CATGTGCGGCTTCACACAGCTTTGTGGACACAGCTGAGAGGCGCACTGAGCCCGCTTACC 8494  
QY 61 GCCCAGAGCCTTTCTTCCAGTGCAGATGCGCCGAGCAGCAGCAGCAGCAGTGGGCTCT 120  
Db 8493 GCCCAGAGCCTTTCTTCCAGTGCAGATGCGCCGAGCAGCAGCAGCAGCAGTGGGCTCT 8434  
QY 121 GAGATACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 180  
Db 8433 GAGATACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 8374  
QY 181 GCCCAGAGGCGGAGAGGCGCTCTCTGTGTGAGTTCAGAGCAGCAGCAGCAGCAGCAG 240  
Db 8373 GCCCAGAGGCGGAGAGGCGCTCTCTGTGTGAGTTCAGAGCAGCAGCAGCAGCAGCAG 8314  
QY 241 GAGATCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 300  
Db 8313 GAGATCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 8254  
QY 301 GAGATCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 360  
Db 8253 GAGATCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 8194  
QY 361 GTACCCCTGCTGTGAGCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 420  
Db 8193 GTACCCCTGCTGTGAGCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 8135  
QY 421 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 479  
Db 8134 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8075  
QY 480 TGGAGCTGATGCGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 539  
Db 8074 TGGAGCTGATGCGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8015  
QY 540 ACCGAGATGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 599  
Db 8014 ACCGAGATGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7955  
QY 600 TGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659  
Db 7954 TGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7895  
QY 660 AGTCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 719  
Db 7894 AGTCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7835  
QY 720 CTGCAATCGATTTCTGAGAGGCTTGGGAGCCTCGAGTGAAGTCCGCTGTGACACAGCAGC 779  
Db 7834 CTGCAATCGATTTCTGAGAGGCTTGGGAGCCTCGAGTGAAGTCCGCTGTGACACAGCAGC 7775  
QY 780 CTTCCGAGACCCGAGATGATGAGAAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 839  
Db 7774 CTTCCGAGACCCGAGATGATGAGAAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 7715



840 TAATCTTGTACTTCAAGAGAGGAATCAAGCTCGAGAAACAATTGAAGCAGAAATTCTGCC 899  
|||||  
7714 TAATCTTGTACTTCAAGAGAGGAATCAAGCTCGAGAAACAATTGAAGCAGAAATTCTGCC 7655  
QY AGCGCTATGACCAAGCTCATGAGAGCCTTGAAAAAAAGGTGAAGCCATCGAAAAACACC 959  
Db |||||  
7654 AACGCTATGACCAAGCTCATGAGAGCCTTGAAAAAAAGGTGAAGCCATCGAAAAACACC 7595  
QY CCGCGCCGGGGGCGCAAGAGAGCAAGTCCGCGAAGTAACTAACAAAAAGCATTTCCCTGAGA 1019  
Db |||||  
7594 CCGCGCGGGGGCGCAAGAGAGCAAGTTCGCGAAGTAACTAACAAAAAGCATTTCCCTGAGA 7535  
QY 1020 TCCGCAAGAGCGCGAGCTGCAAGAGCGCATGCAAGCAGAGGTGAGCCAGCGGGCAGTG 1079  
Db |||||  
7534 TCCGCAAGAGCGCGAGCTGCAAGAGCGCATGAG---AGGGTGGGCGCAGCGGGGCGAGTG 7478  
QY 1080 GGCTGTCCATGTGCGCGCGCCGCGAGAGCAAGAGGTTCAGAGATCATGATGAGCTCTT 1139  
Db |||||  
7477 GGCTGTCCATGTGCGCGCGCCGCGAGAGCAAGAGGTTCAGAGATCATGATGAGCTCTT 7418  
QY 1140 CAGAGCAGAGAAACCTGAGAAAGAGATGCGCGACGCTGCGCCGATCCCGCCCATGCTGT 1199  
Db |||||  
7417 CAGAGCAGAGAAACCTGAGAAAGCAGATGCGCGACGCTGCGCCGATCCCGCCCATGCTGT 7358  
QY 1200 AGCAGCTGACAGCAGCGCATCAAGTTGATCAACATGAACGGGCTTATGCGCACCCCA 1259  
Db 7357 AGCAGCTGACAGCAGCGCATCAAGTTGATCAACATGAACGGGCTTATGCGCACCCCA 7298  
QY 1260 TGAAGGTGTACAAAGACCGCCAGGTGATGAACAATGTGAATGAGCAGAGAAAGAGACCT 1319  
Db |||||  
7297 TGAAGGTGTACAAAGACCGCCAGGTGATGAACAATGTGAATGAGCAGAGAAAGAGACCT 7238  
QY 1320 TCCGGGAGAAAGTTGATGAGCAGATCCCAAGAACTTTGGCCTGATCGATCTTCTTGAAGA 1379  
Db |||||  
7237 TCCGGGAGAAAGTTGATGAGCAGATCCCAAGAACTTTGGCCTGATCGATCTTCTTGAAGA 7178  
QY 1380 GGAAGAAGTGTGAGTGAAGTGCCTCTATTTACTGACTGAAGTGAAGTGAAGTATATA 1439  
Db 7177 GGAAGAAGTGTGAGTGAAGTGCCTCTATTTACTGACTGAAGTGAAGTGAAGTATATA 7118  
QY 1440 AGAGCTGTGTGAGACGAGCTATTCGCGCGCGGCAAGAGCCAGCAGCAACAACAGCAGC 1499  
Db 7117 AGAGCTGTGTGAGACGAGCTATTCGCGCGCGGCAAGAAACCAAGCAGCAACAACAAGC 7058  
QY 1500 AGCAGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1559  
Db |||||  
7057 CGGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7000  
QY 1560 AAGATGAGAAAGAGAGAGAAAGAGAGCGGAGAGAGAGAGAGAGCGGAGGTGAGAGA 1619  
Db 6999 AAGATGAGAAAGAGAGAGAAAGAGAGCGGAGAGAGAGAGAGAGCGGAGGTGAGAGA 6940  
QY 1620 AGCAGCAGAGAGAGCCTCTCAAGAGAGAGAGAGAGCAGCAGCCTCAAGAGAGAGAGAGAG 1679  
Db 6939 AGCAGCAGAGAGAGCCTCTCAAGAGAGAGAGAGAGCAGCAGCCTCAAGAGAGAGAGAG 6880  
QY 1680 AGAAGAGAGGCTGTGCTCTCAAAAGGCGCAAACTGCGCAACGCGGAGAGAGAGAGAG 1739  
Db 6879 AGAAGAGAGGCTGTGCTCTCAAAAGGCGCAAACTGCGCAACGCGGAGAGAGAGAGAG 6820  
QY 1740 GCGCGATCAACCGGCTCAATGAGCTTAATGAGGCGCAACGAGAGAGAGGAGCATCAACCCCGCAG 1799  
Db 6819 GCGCGATCAACCGGCTCAATGAGCTTAATGAGGCGCAACGAGAGAGAGGAGCATCAACCCCGCAG 6760  
QY 1800 AGAGCGCCAGAGTGGCTCTCAAGAGAGAGAGAGAGAGTTCGCTGAGAGAGAGAGAGAG 1859  
Db 6759 AGAGCGCCAGAGTGGCTCTCAAGAGAGAGAGAGAGTTCGCTGAGAGAGAGAGAGAG 6700  
QY 1860 TGGAAAAGAGCAGAGAAAGGTCTCTTGAAGACAGCGCGCGCACTGCTGAGAGAGAGAG 1919  
Db 6699 TGGAAAAGAGCAGAGAAAGGTCTCTTGAAGACAGCGCGCGCACTGCTGAGAGAGAGAG 6640  
QY 1920 TGGTGGGCTCAAGAGTGTGTGCAAGTGTAAAGAACTTCACTCAACAGAGAGAGCG 1979

6639 TGGTGGGCTCAAGAGTGTGTGCAAGTGTAAAGAACTTCACTCAACAGAGAGAGCG 6580  
Db |||||  
1980 AGAAGCTGAGAGAGATCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2039  
QY |||||  
6579 AGAAGCTGATGAGATCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6520  
Db |||||  
2040 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2099  
QY |||||  
6519 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6460  
Db |||||  
2100 TGAAGATGAGAGAGATGAGAGGCGTGGGCGTGAAGCGGAAATGAGAGAGAGATGCTGAAG 2159  
QY 6459 TGAAGATGAGAGAGATGAGAGGCGTGGGCGTGAAGCGGAAATGAGAGAGAGATGCTGAAG 6400  
Db |||||  
2160 AGGCTGAAGCTTCAATGCTCTGGGAAATGAGAGTCCCAAGAGGAGATCAAGTGGCCCA 2219  
QY |||||  
6399 AGGCTGA-----AG 6391  
Db |||||  
2220 CCACTGTCAACAAGAGCTGAGACACCGAGAGAGATCCCTCTCTCAACTGAGAGCGCGCA 2279  
QY |||||  
6390 CCACTGTCAACAAGAGCTGAGACACCGAGAGAGATCCCTCTCTCAACTGAGAGCGCGCA 6331  
Db |||||  
2280 AGAGACAGAGGAGAGATGGGCGCCAAAGCGCCAGCCACCTGTGGGCGCGACCGGCGCACCC 2339  
QY 6330 AGAGACAGAGGAGAGATGGGCGCCAAAGCGCCAGCCACCTGTGGGCGCGACCGGCGCACCC 6271  
Db |||||  
2340 CAGGCGCACCCACCCCAACCGAGAGACATCCCGGGCGCCCATTTGAGGCCACCCCGGCGCT 2399  
QY 6270 CAGGCGCACCCACCCCAACCGAGAGACATCCCGGGCGCCCATTTGAGGTCAACCCCGGCGCT 6211  
Db |||||  
2400 CTGAAGCCACCGGAGCGCTTACCGCCCGCACAGACCCCGCATTCGCTCTGCACTCTCTC 2459  
QY 6210 CTGAAGCCACCTTACCGCTTACCGCCCGCACAGACCCCGCATTCGCTCTTCACTCTCTC 6151  
Db |||||  
2460 CTGTGTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2519  
QY 6150 CTGTGTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6091  
Db |||||  
2520 GGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2579  
QY 6090 GGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6031  
Db |||||  
2580 AGCGCGTCAAGAGCGAGGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2639  
QY 6030 AGCGCGTCAAGAGCGAGGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5971  
Db |||||  
2640 CGAGGCGCGTGAAGGCAAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2699  
QY 5970 CGAGGCGCGTGAAGGCAAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5911  
Db |||||  
2700 GCGGAGAGGCGCACACTGCGCAAGAGCTGGGGCGCGCCCAAGAGAGAGAGAGAGAGAG 2759  
QY 5910 GCGGAGAGGCGCACACTGCGCAAGAGCTGGGGCGCGCCCAAGAGAGAGAGAGAGAGAG 5851  
Db |||||  
2760 CCGTGAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2819  
QY 5850 CCGTGAAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5791  
Db |||||  
2820 CAAAGGCCAGCTCTCAACCGGAGCTGGGAGACCCCGGAGCAATGCTCAACCCCAAGAGC 2879  
QY 5790 CAAAGGCCAGCTCTCAACCGGAGCTGGGAGACCCCGGAGCAATGCTCAACCCCAAGAGC 5731  
Db |||||  
2880 CACTGAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2939  
QY 5730 CACTGAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5671  
Db |||||  
2940 AAGTCCATGAGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2999  
QY 5670 AAGTCCATGAGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5611  
Db |||||  
3000 CGCAGCCGAGAAACCTGCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3059  
QY |||||



Db 5610 GCACACCCGAAAACTGACGCCGAGAGCAGCCCTTCAGCAGCTGCGACAGACCCCC 5551  
Oy 3060 GAGGACAAGACGAGAGCCCGGACCCCGCCGACCAAGAG----- 3100  
Db 5550 GGGGACAAGACGAGAGCCCGGACCCCGCCGACCAAGAGGACAGAGAGCTGTGTTCT 5491  
Oy 3101 -----GCTTGGAGAGCCGACCAAGAGCTGCTGGGAGCCCTTGTGAGACTTCG 3155  
Db 5490 TCCAGCTTGGCAGCCGAGGCCAGAGAGCTGCTGGGAGCCCTTGTGAGACTTCG 5431  
Oy 3156 GCTGCCCCCTGCCCCGCCCCCGGAGAGATCAAGGCTCCCCGAGATGCCCCGAGAC 3215  
Db 5430 GCTGCCCCCTGCCCCGCCCCCGGAGAGATCAAGGCTCCCCGAGATGCCCCGAGAC 5371  
Oy 3216 CCTCAGCTTCTCTACGCTCACTGTGACCCACTGCCCCCTGGGCTTCATGACCTG 3275  
Db 5370 CCTCAGCTTCTCTACGCTCACTGTGACCCACTGCCCCCTGGGCTTCATGACCTG 5311  
Oy 3276 CCGGCGCGTCTGCGCGGCCACCCACCATCTCAACCCGCTTCCTCATCTCTCTG 3335  
Db 5310 CCGGCGCGTCTGCGCGGCCACCCACCATCTCAACCCGCTTCCTCATCTCTCTG 5251  
Oy 3336 CCAGACACCCGACGCTCTCCGAGAGCAAAATAGTGCCATCTCCAGAGAAATGTGCTC 3395  
Db 5250 CCAGACACCCGACGCTCTCCGAGAGCAAAATAGTGCCATCTCCAGAGAAATGTGCTC 5191  
Oy 3396 AGCTCAGCTCCCGTACTCAGAGCATGCAAGGACCCGGTGGGCTCTGACCATGGGC 3455  
Db 5190 AGCTCAGCTCCCGTACTCAGAGCATGCAAGGACCCGGTGGGCTCTGACCATGGGC 5131  
Oy 3456 TGCCCCCTCCATGAGACCCCAAAAAGCTGGCACCTTCAGCGAGTGAAGCAGACAGC 3515  
Db 5130 TGCCCCCTCCATGAGACCCCAAAAAGCTGGCACCTTCAGCGAGTGAAGCAGACAGC 5071  
Oy 3516 TGTCGCCACGAGGAGCCAGGCTGGGCCACCGGAGAGCTGGGGTGGCCACAGCCGAGAG 3575  
Db 5070 TGTCGCCACGAGGAGCCAGGCTGGGCCACCGGAGAGCTGGGGTGGCCACAGCCGAGAG 5011  
Oy 3576 CGTCCGCTCTGAGAGGAGCAGCTCTGGGCTCAGTTCGGGCGGAGAGCATCAACAAGCA 3635  
Db 5010 CGTCCGCTCTGAGAGGAGCAGCTCTGGGCTCAGTTCGGGCGGAGAGCATCAACAAGCA 4951  
Oy 3636 TTCCGAGCACAAGGATGCCCTTCGGAACAGGCCATCATACCGGCTTCATCACCCAG 3695  
Db 4950 TTCCGAGCACAAGGATGCCCTTCGGAACAGGCCATCATACCGGCTTCATCACCCAG 4891  
Oy 3696 GCAAGCCGAGCTGACGCTCTGTATCAAGGGCAGCATCAACGAGATCATTCGGCGAGAGCAGCC 3755  
Db 4890 GCAAGCCGAGCTGACGCTCTGTATCAAGGGCAGCATCAACGAGATCATTCGGCGAGAGCAGCC 4831  
Oy 3756 CGAGTCGCTTGGACCGCGGCGCGGAGAGCAGGCTGCCAAGGGGACCGTCATCTACGAG 3815  
Db 4830 CGAGTCGCTTGGACCGCGGCGCGGAGAGCAGGCTGCCAAGGGGACCGTCATCTACGAG 4771  
Oy 3816 GCAAGAGGGCCACGCTCTTGTCTATAGAGGTGAGATGCTGTGACCCAGTGTCCAAAG 3875  
Db 4770 GCAAGAGGGCCACGCTCTTGTCTATAGAGGTGAGATGCTGTGACCCAGTGTCCAAAG 4711  
Oy 3876 AGGAGCGGAGAGAGCTCAAGAGACCCCGCATGAGACGGCGCGCCCAAGCGGACCTATG 3935  
Db 4710 AGGAGCGGAGAGAGCTCAAGAGACCCCGCATGAGACGGCGCGCCCAAGCGGACCTATG 4651  
Oy 3936 ACATGATGAGGGGCGCGTGGGACAGGCAATCTCCAGCCAGCATCGAAGGTCTCATGG 3995  
Db 4650 ACATGATGAGGGGCGCGTGGGACAGGCAATCTCCAGCCAGCATCGAAGGTCTCATGG 4591  
Oy 3996 GCGGTGCATCCCGCGGAGCGACACAGCCCAACCACTCAAGAGCAGCAGCATCC 4055  
Db 4590 GCGGTGCATCCCGCGGAGCGACACAGCCCAACCACTCAAGAGCAGCAGCATCC 4531  
Oy 4056 GCGGTGCATCAACAAGAGATCCTCGGTCTTACGTGAGGCAAGAGAGCTACCTGC 4115  
Db 4530 GCGGTGCATCAACAAGAGATCCTCGGTCTTACGTGAGGCAAGAGAGCTACCTGC 4471

Oy 4116 GTGAGAGCCAGAGCTCTTAAAGCGGAGAGGACAGCTTCGCCCCACCGCCTTCAAGG 4175  
Db 4470 GTGAGAGCCAGAGCTCTTAAAGCGGAGAGGACAGCTTCGCCCCACCGCCTTCAAGG 4411  
Oy 4176 ACTGACCGAGGCTTCAAGAGCAGAGGCTTGGGCTCCCTGAAGCTGAAGCTGGCCATG 4235  
Db 4410 ACTGACCGAGGCTTCAAGAGCAGAGGCTTGGGCTCCCTGAAGCTGAAGCTGGCCATG 4351  
Oy 4236 AGGAGCTGTGGCCAGGAGGAGAGGAGGCGGCTCCATCATGAGATCCGCGGAGG 4295  
Db 4350 AGGAGCTGTGGCCAGGAGGAGAGGAGGCGGCTCCATCATGAGATCCGCGGAGG 4291  
Oy 4296 AGCTGCGGACACGCGCGAGGCTCCCTGAGCCCGCGCGGCTCAAGAGAGGCTCATCA 4355  
Db 4290 AGCTGCGGACACGCGCGAGGCTCCCTGAGGCGGCTGAGGAGGAGGCTCATCA 4231  
Oy 4356 CGCAGGAGCAGCCCGCTCAAGTACGACACCGGCGCTCAACCTGGCTTCAAAAAGCAG 4415  
Db 4230 CGCAGGAGCAGCCCGCTCAAGTACGACACCGGCGCTTCAACCTGGCTTCAAAAAGCAG 4171  
Oy 4416 AGTACGCTCCCTCATGCGGAGGCGCGCGGAGGTTCCACCGGTGACCCGCTGATG 4475  
Db 4170 AGTACGCTCCCTCATGCGGAGGCGCGCGGAGGTTCCACCGGTGACCCGCTGATG 4111  
Oy 4476 TGATGCGGACGCGCGGACCTGGAACGTGCTGCTACGAGAGAGCTGAGAGCCGAGC 4535  
Db 4110 TGATGCGGACGCGCGGACCTGGAACGTGCTGCTACGAGAGAGCTGAGAGCCGAGC 4051  
Oy 4536 CAGGAGCCGCAAGCAGCTCGGGGGGCTCATTTGCGCGGCGGCCCTCATTTGCTG 4595  
Db 4050 CAGGAGCCGCAAGCAGCTCGGGGGGCTCATTTGCGCGGCGGCCCTCATTTGCTG 3991  
Oy 4596 AGCTGGTAAAGCGGAGAGGCGCCCTGACCTATGAGAGCAACAGGGGACCTTTGCGG 4655  
Db 3990 AGCTGGTAAAGCGGAGAGGCGCCCTGACCTATGAGAGCAACAGGGGACCTTTGCGG 3931  
Oy 4656 GCAACCTCCACAGAGGTTGCGCCGCTGACCATGCGGAGACCCACGCGGCTTCAGAGG 4715  
Db 3930 GCAACCTCCACAGAGGTTGCGCCGCTGACCATGCGGAGACCCACGCGGCTTCAGAGG 3871  
Oy 4716 GCAAGCTTTGCTCAGCAAGGATCCAGAGCCGAAGCTGACGTGACGCTGTGAGA 4775  
Db 3870 GCAAGCTTTGCTCAGCAAGGATCCAGAGCCGAAGCTGACGTGACGCTGTGAGA 3811  
Oy 4776 TGCGCAAGTCCCGGACAGACCGGCGCGGAGACCAACCAACCCATTCGCGCTATG 4835  
Db 3810 TGCGCAAGTCCCGGACAGACCGGCGCGGAGACCAACCAACCCATTCGCGCTATG 3751  
Oy 4836 AGCAGCTGCTTCGAGGCGTGAAGTGGCTGTGACTGTATCGAGGCCACATCCCGTGGCT 4895  
Db 3750 AGCAGCTGCTTCGAGGCGTGAAGTGGCTGTGACTGTATCGAGGCCACATCCCGTGGCT 3691  
Oy 4896 TGACCCCACTCCATACCCCGGAGATCCCTCTGAGCGAGCGGTGCTACTACCTGC 4955  
Db 3690 TGACCCCACTCCATACCCCGGAGATCCCTCTGAGCGAGCGGTGCTACTACCTGC 3631  
Oy 4956 CCGGACACTGGGCCCCCAACCCCACTACCCGACCTGTAACCAACCTACCTACCGGG 5015  
Db 3630 CCGGACACTGGGCCCCCAACCCCACTACCCGACCTGTAACCAACCTACCTACCGGG 3571  
Oy 5016 GCTACCCGACAGCGCGGCGCTGAGAAACCGGAGACCATCATCAATGACTACACT 5075  
Db 3570 GCTACCCGACAGCGCGGCGCTGAGAAACCGGAGACCATCATCAATGACTACACT 3511  
Oy 5076 CGCAGCAGATGACCAACAACGCGCACCGCATGGCCCAAGAGCTGATGCTGAGGG 5135  
Db 3510 CGCAGCAGATGACCAACAACGCGCACCGCATGGCCCAAGAGCTGATGCTGAGGG 3451  
Oy 5136 GCTTGTGCGCCCGGAGGCTTCGCTGAGCACTCAATTAGCTGCGGGTCCCGAGGACATA 5195  
Db 3450 GCTTGTGCGCCCGGAGGCTTCGCTGAGCACTCAATTAGCTGCGGGTCCCGAGGACATA 3391

[illegible]

Dp	2310	GCCTCGTGAAGCTTGGCGGGGAGGCGCCCACTTCCACACTTCGGGCTCGTGGCTGAGA	2251
Qy	6336	GCCAGCCCTCGTCCAGGCCCGCTGCTCCAGACCCGCCCGAGGGGTCAAGGTCAACGAGCGGG	6395
Dp	2250	GCCAGCCCTCGTCCAGGCCCGCTGCTCCAGACCCGCCCGAGGGGTCAAGGTCAACGAGCGGG	2191
Qy	6396	TGGTCACTCCCTGGCCAGACATCATGTAGGTCAATCAACAGAACTAACCCCGGACCAACC	6455
Dp	2190	TGGTCACTCCCTGGCCAGACATCATGTAGGTCAATCAACAGAACTAACCCCGGACCAACC	2131
Qy	6456	CACAGCAAGCTCAGGGGACCCCTGGCCGCGCCCTCTACTCTCTTCCCGGGGCGAGCTGCC	6515
Dp	2130	CACAGCAAGCTCAGGGGACCCCTGGCCGCGCCCTCTACTCTCTTCCCGGGGCGAGCTGCC	2071
Qy	6516	CCGACTGTGACCTTCCGCGGCCCAACCAAGTAACTCTACCTCCGCGCCCGGACCATGGTG	6575
Dp	2070	CCGACTGTGACCTTCCGCGGCCCAACCAAGTAACTCTACCTCCGCGCCCGGACCATGGTG	2011
Qy	6576	CCCCGGCCCGTGGCTCCCCCAACAGCGAAGGGGCAAGAGGTCTCCAGAGCCAAACAAGA	6635
Dp	2010	CCCCGGCCCGTGGCTCCCCCAACAGCGAAGGGGCAAGAGGTCTCCAGAGCCAAACAAGA	1951
Qy	6636	CGTGGCTCTTGGGTGGTGGTGAAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGA	6695
Dp	1950	CGTGGCTCTTGGGTGGTGGTGAAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGA	1891
Qy	6696	CGGAGCCAGGGCACTCCCGGAGTCTGTGTAAACCGCTGCTGTACCGGATATGGGAAACAGA	6755
Dp	1890	CGGAGCCAGGGCACTCCCGGAGTCTGTGTAAACCGCTGCTGTACCGGATATGGGAAACAGA	1831
Qy	6756	CGGAGCCAGAGATGGGCTCCAAAGTCTCCAGGCAACACAGCCAGCCGCGACCTTCT	6815
Dp	1830	CGGAGCCAGAGATGGGCTCCAAAGTCTCCAGGCAACACAGCCAGCCGCGACCTTCT	1771
Qy	6816	TCAGCAAGCTGACCGGAGAGCACTCCGCGCATGTCTAAAGTCCAAAGAACAGAGATTCAACA	6875
Dp	1770	TCAGCAAGCTGACCGGAGAGCACTCCGCGCATGTCTAAAGTCCAAAGAACAGAGATTCAACA	1711
Qy	6876	AGAAAGTGAACACCCCAACACCGGAATAGGCTGAATACAAATATACGCGAGCTTGGAGCGG	6935
Dp	1710	AGAAAGTGAACACCCCAACACCGGAATAGGCTGAATACAAATATACGCGAGCTTGGAGCGG	1651
Qy	6936	AGATCTTCAATATATGCCCGCATCAACCGGAAACAGGCTTATATGACCTATATGAAGCCAGGCGG	6995
Dp	1650	AGATCTTCAATATATGCCCGCATCAACCGGAAACAGGCTTATATGACCTATATGAAGCCAGGCGG	1591
Qy	6996	TGCAGGAACATGCGACGACCAACATGGGGGCTGGAGGGCATATATTAAGAAAGGCACTCATGG	7055
Dp	1590	TGCAGGAACATGCGACGACCAACATGGGGGCTGGAGGGCATATATTAAGAAAGGCACTCATGG	1531
Qy	7056	GTAAATATGACAGATGGGAAGAGTCCCCCGCGCTCAAGGCGCAATGCTTTTAACCCCTCTGA	7115
Dp	1530	GTAAATATGACAGATGGGAAGAGTCCCCCGCGCTCAAGGCGCAATGCTTTTAACCCCTCTGA	1471
Qy	7116	ATGCACTGACAGGCTCCCGCTGCTATATGCCCATAACCGCTGCTGACGAGCGAGTGAACC	7175
Dp	1470	ATGCACTGACAGGCTCCCGCTGCTATATGCCCATAACCGCTGCTGACGAGCGAGTGAACC	1411
Qy	7176	ACAACACTCACCTCGCCAGGTGGCGGGGAAAGGCCAAGGTCTCTGGAGACCCACGACGCC	7235
Dp	1410	ACAACACTCACCTCGCCAGGTGGCGGGGAAAGGCCAAGGTCTCTGGAGACCCACGACGCC	1351
Qy	7236	GAAAAAGCAAGTCCCGGGCCCCCGGCTTGGCATTTGGGGACCGGCAACCTCTGTCTCTCT	7295
Dp	1350	GAAAAAGCAAGTCCCGGGCCCCCGGCTTGGCATTTGGGGACCGGCAACCTCTGTCTCTCT	1291
Qy	7296	CAGTGCACCTCGAGAGGAGACTGCAACCGCGCGGACCGCGCTCAACACCGCGTGTGGGAGG	7355
Dp	1290	CAGTGCACCTCGAGAGGAGACTGCAACCGCGCGGACCGCGCTCAACACCGCGTGTGGGAGG	1231
Qy	7356	ACAGGCCCCTGCTCCGACAGTTTCAACGCGCATTTCCCTTACAAACCCCTGATCATATCGGCTGC	7415



Db 541 CGAAGATCAACATGTGTAGAGCAGACATCTCTTAAGCTGAAAGAGCAACAGACTG 600  
OY 602 GAGGAGGAGGCTGCGCAAGCGCCGAGCCTGAGAAAGCCGCTGTCAACGCGGCCATCGAG 661  
Db 601 GAGGAGGAGGCTGCGCAAGCGCCGAGCCTGAGAAACCGGTGTACCGCGGCCATCGAG 660  
OY 662 TCGAAGCACCGGAGCCTGTGTGAGATCATCTACGAGAGAACCGAAGAAAGCTGAGACT 721  
Db 661 TCGAAGCACCGGAGCCTGTGTGAGATCATCTACGAGAGAACCGAAGAAAGCTGAGACT 720  
OY 722 GCAATCGGATCTGTGAAGGCTGTGGGCCCCAGGTGAGCTGCGCTGTACAAACGAGCC 781  
Db 721 GCAATCGGATCTGTGAAGGCTGTGGGCCCCAGGTGAGCTGCGCTGTACAAACGAGCC 780  
OY 782 TCCGACACCCGGAGATATGAGAACATCAAAATTAACGAGCGATGTGGAAAGAGCTTA 841  
Db 781 TCCGACACCCGGAGATATGAGAACATCAAAATTAACGAGCGATGTGGAAAGAGCTTA 840  
OY 842 ATCTTGATCTTCAAGAGAGAAATCAGCTCGGAAACAATGAGAGAGAACTTCTGACAG 901  
Db 841 ATCTTGATCTTCAAGAGAGAAATCAGCTCGGAAACAATGAGAGAGAACTTCTGACAG 900  
OY 902 CGCTATGACAGACTCATGAGGCTTTGAAAAAAAGGTGAGCGCATGAAAAACAACCCG 961  
Db 901 CGCTATGACAGACTCATGAGGCTTTGAAAAAAAGGTGAGCGCATGAAAAACAACCCG 960  
OY 962 CGCGCGCGGCGCAAGAGAGCAAGGTGCGCGAGTACTAGAAAAAGAGTTCCCTGAGATC 1021  
Db 961 CGCGCGCGGCGCAAGAGAGCAAGGTGCGCGAGTACTAGAAAAAGAGTTCCCTGAGATC 1020  
OY 1022 CGCAAGCACGCGAGCTGTGAGAGCGCATGACAGAGGCTGGGCGAGCGGGCACTGGG 1081  
Db 1021 CGCAAGCACGCGAGCTGTGAGAGCGCATGACAGAGGCTGGGCGAGCGGGCACTGGG 1080  
OY 1082 CTGTCTATGTGCGCGCGCGCGAGCGAGCAAGAGTGTGAGATCATGATGAGCTTCTCA 1141  
Db 1081 CTGTCTATGTGCGCGCGCGCGAGCGAGCAAGAGTGTGAGATCATGATGAGCTTCTCA 1140  
OY 1142 GAGCAGAGAACTCTGAGAGAGAGATGCGCGAGCTGGCCGTGATCCCGGCCATGCTGTAC 1201  
Db 1141 GAGCAGAGAACTCTGAGAGAGAGATGCGCGAGCTGGCCGTGATCCCGGCCATGCTGTAC 1200  
OY 1202 GACGCTGACGAGGCGCATCAAGTTTCATCAACATGAAGCGGCTTATGGCCGACCCCATG 1261  
Db 1201 GACGCTGACGAGGCGCATCAAGTTTCATCAACATGAAGCGGCTTATGGCCGACCCCATG 1260  
OY 1262 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGTGAGTGAAGCAGAGAGAGACCTTC 1321  
Db 1261 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGTGAGTGAAGAGAGAGACCTTC 1320  
OY 1322 CGGAGAGAGTTCAATGACATATCCCAAGAACTTTGGCTGTATGSCATCATTTCTGAGAG 1381  
Db 1321 CGGAGAGAGTTCAATGACATATCCCAAGAACTTTGGCTGTATGSCATCATTTCTGAGAG 1380  
OY 1382 AAGACAGTGGCTGAGTGGCTCTTATTACTTACTACCTTAAGAAAGTGAAGACTTAAG 1441  
Db 1381 AAGACAGTGGCTGAGTGGCTCTTATTACTTACTACCTTAAGAAAGTGAAGACTTAAG 1440  
OY 1442 AAGCTGTGTGAGACGAGACTATCGCGCGCGGCAAGAGCCAGCAGCAACAACGACAG 1501  
Db 1441 AAGCTGTGTGAGACGAGACTATCGCGCGCGGCAAGAGCCAGCAGCAACAACGACAG 1500  
OY 1502 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCTTCCCGCAGCAGCCAGAGAGAGAA 1561  
Db 1501 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCTTCCCGCAGCAGCCAGAGAGAGAA 1560  
OY 1562 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621  
Db 1561 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
OY 1622 GACAAAGAGAGAGCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681  
Db 1621 GACAAAGAGAGAGCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680

OY 1682 AAGAGGCTGTGGCTTCAAAAGCGCAAAACTGGCCAAAGAGGAGAGAGAGAGAGAGAG 1741  
Db 1681 AAGAGGCTGTGGCTTCAAAAGCGCAAAACTGGCCAAAGAGGAGAGAGAGAGAGAGAG 1740  
OY 1742 CGCATCACCGGCTCAATGGCTTAATGAGGCGCAACAGGAGAGGCTCATCCCCAGAG 1801  
Db 1741 CGCATCACCGGCTCAATGGCTTAATGAGGCGCAACAGGAGAGGCTCATCCCCAGAG 1800  
OY 1802 AGCGCGAGCTGGCTCCATGAGAGCTGATGAGAGTTCTCGCTGACAGAGAGAGAAATG 1861  
Db 1801 AGCGCGAGCTGGCTCCATGAGAGCTGATGAGAGTTCTCGCTGACAGAGAGAGAAATG 1860  
OY 1862 GAAACAGCAAGAAAGGCTCTCTGGAACAGGCGCGCAACTGTGCGCAATGCGCCGAGATG 1921  
Db 1861 GAAACAGCAAGAAAGGCTCTCTGGAACAGGCGCGCAACTGTGCGCAATGCGCCGAGATG 1920  
OY 1922 GTGGGCTCCAGACTGTGTGCGAGTGTAAAGAACTTCTTACTTCACTTAAGAGAGCAG 1981  
Db 1921 GTGGGCTCCAGACTGTGTGCGAGTGTAAAGAACTTCTTACTTCACTTAAGAGAGCAG 1980  
OY 1982 AACCTGATGAGATCTTGTGAGAGCACAAGCTGAAATGAGAGAGAGAGAGAGAGAG 2041  
Db 1981 AACCTGATGAGATCTTGTGAGAGCACAAGCTGAAATGAGAGAGAGAGAGAGAGAG 2040  
OY 2042 AGGAGAGAGAGAGAGAGAGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 2101  
Db 2041 AGGAGAGAGAGAGAGAGAGCGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
OY 2102 GAGGATGAGAGAGATGAGAGCGGTGCGGCGGTGAGCGGAAATGAGAGAGAGATGTGTAGAGAG 2161  
Db 2101 GAGGATGAGAGAGATGAGAGCGGTGCGGCGGTGAGCGGAAATGAGAGAGAGATGTGTAGAGAGAG 2160  
OY 2162 GCTGAAGCTTTCATATCTCTCTGTGGAATGAGGTGCCAGAGGGGAAATGCAAGTGGCCCAAGC 2221  
Db 2161 GCTGAAGCTTTCATATCTCTCTGTGGAATGAGGTGCCAGAGGGGAAATGCAAGTGGCCCAAGC 2220  
OY 2222 ACTGTCAACAAAGCTCAGACACCGAGACATCCCTCTCTCACACTGAGAGCGGCAAG 2281  
Db 2221 ACTGTCAACAAAGCTCAGACACCGAGACATCCCTCTCTCACACTGAGAGCGGCAAG 2280  
OY 2282 GACACAGGCGAGATGAGGCGCAAGCGCCGAGCAACCTGAGCGCGAGCGGCAACCCCA 2341  
Db 2281 GACACAGGCGAGATGAGGCGCAAGCGCCGAGCAACCTGAGCGCGAGCGGCAACCCCA 2340  
OY 2342 GGGCCACCCACCCACACGAGAGAGACATCCCGGAGCCCAATTGAGGCCACCCGGGCTCT 2401  
Db 2341 GGGCCACCCACCCACACGAGAGAGACATCCCGGAGCCCAATTGAGGCCACCCGGGCTCT 2400  
OY 2402 GAAAGCCAGGAGCGGCTTACGCGCCCAAGCAGCAACCCCAATGCGGCTTGTGACCTCTCGCT 2461  
Db 2401 GAAAGCCAGGAGCGGCTTACGCGCCCAAGCAGCAACCCCAATGCGGCTTGTGACCTCTCGCT 2460  
OY 2462 GTGTGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2521  
Db 2461 GTGTGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
OY 2522 GAGAGACAGAGAGCGCCCGCGCTGAGAGCTGTGAGCTGTGACACAGAGAGAGAGAGAGAG 2581  
Db 2521 GAGAGACAGAGAGCGCCCGCGCTGAGAGCTGTGAGCTGTGACACAGAGAGAGAGAGAGAG 2580  
OY 2582 CCGGTCAAGAGAGAGTGTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2641  
Db 2581 CCGGTCAAGAGAGAGTGTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
OY 2642 GAGCGCGCTGAGAGCACAAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2701  
Db 2641 GAGCGCGCTGAGAGCACAAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
OY 2702 GGCAGAGGCAACACTGTCCAAAGAGCTTGGGCGCCCTCCAGAGACAGCGACTTCAGTGTCTAAC 2761  
Db 2701 GGCAGAGGCAACACTGTCCAAAGAGCTTGGGCGCCCTCCAGAGACAGCGACTTCAGTGTCTAAC 2760

OY	2762	TGCAGTGCAGACAGAGTGAATGAGGCGCGAGGGCGGCGACAAGAAACGGGTGTGTGCCCA	2812
Db	2761	TGCATGTGCAGACGAGGTGTATGAGCCCGAGGGCGGCGACAGAAACGGGTGTGTGCCCA	2820
OY	2822	AGGCCCAAGCCTTCTCAACCCCGACCTTGCGACACCCCGGGGCAATGTCTCAACCCCAAGAGCCA	2881
Db	2821	AGGCCCAAGCCTTCTCAACCCCGACCTTGCGACACCCCGGGGCAATGTCTCAACCCCAAGAGCCA	2880
OY	2882	CTGGAACCTGAAGCAGCTGTAAAGCAGCGAGCGCTGTGCATCCCTCCCATCAAGTCAACAA	2941
Db	2881	CTGGACCTGAAGCAGCTGTAAAGCAGCGAGCGGTGTGCATCCCTCCCATCAAGTCAACAA	2940
OY	2942	GTCCATGAGCCCCCGGGAGGAGCGCAGCTCCCAACAAGCAGCTCCCGACGCCCAACG	3001
Db	2941	GTCCATGAGCCCCCGGGAGGAGCGCAGCTCCCAACAAGCAGCTCCCGACGCCCAACG	3000
OY	3002	CCACCGCAAAACCTGCAGCCCGAGAGCAGCGCCCTCAGCAGCCTGGCAGAGCGCCCGG	3061
Db	3001	CCACCGCAAAACCTGCAGCCCGAGAGCAGCGCCCTCAGCAGCCTGGCAGAGCGCCCGG	3060
OY	3062	GGCAAGAGCAGAGAGCGCCGGCACACCCCGCGCAACAAGAGGCTTTCGACGCCGAGGCCAG	3121
Db	3061	GGCAAGAGCAGAGAGCGCCGGCACACCCCGCGCAACAAGAGGCTTTCGACGCCGAGGCCAG	3120
OY	3122	AAAGCTGCGTGGGGAGACCCCGCTTGTCTGTGAACCTTCGGGCTGGCCCTTCCCGGTGCCCCCGT	3181
Db	3121	AAAGCTGCGTGGGGAGACCCCGCTTGTCTGTGAACCTTCGGGCTGGCCCTTCCCGGTGCCCCCGT	3180
OY	3182	GAGGTGATCAAGAGCCTCCCGCATGCGCCGAGACCCCTCAGCCTTCTCTTGAAGCTCAACT	3241
Db	3181	GAGGTGATCAAGAGCCTCCCGCATGCGCCGAGACCCCTCAGCCTTCTCTTGAAGCTCAACT	3240
OY	3242	GGTCAACCCAGCTGCCCCCTTGAGGCTTCATGACAATGCGCGGACCGCTGTGTCGCGGCCAACCC	3301
Db	3241	GGTCAACCCAGCTGCCCCCTTGAGGCTTCATGACAATGCGCGGACCGCTGTGTCGCGGCCAACCC	3300
OY	3302	ACCATATTCGAACCCCGGCTCCCTCTCATCTCTCTTGCGAAGACAACCCAGCGTCTCTGAGAGG	3361
Db	3301	ACCATATTCGAACCCCGGCTCCCTCTCATCTCTCTTGCGAAGACAACCCAGCGTCTCTGAGAGG	3360
OY	3362	CAAAATGAGTGCATATCCCAAGAAATGTCCGGTCCAGCTCCAGCTCCCGCTACTCAGAGCAT	3421
Db	3361	CAAAATGAGTGCATATCCCAAGAAATGTCCGGTCCAGCTCCAGCTCCCGCTACTCAGAGCAT	3420
OY	3422	GCCAAAGGCCCCGGTGGGCGCTGTCCATCAATGGGGGCTGCCCTTGCCATGGACCCCAAAAAG	3481
Db	3421	GCCAAAGGCCCCGGTGGGCGCTGTCCATCAATGGGGGCTGCCCTTGCCATGGACCCCAAAAAG	3480
OY	3482	CTGGCAGCCTTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCAACGGGGCGAGGCTTGAGGCCA	3541
Db	3481	CTGGCAGCCTTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCAACGGGGCGAGGCTTGAGGCCA	3540
OY	3542	CCGGAAGGCTCGGGGGTGCACACGCGCCAGGAGGCGTCCGGTGAAGAGGAGCAGCTGTG	3601
Db	3541	CCGGAAGGCTCGGGGGTGCACACGCGCCAGGAGGCGTCCGGTGAAGAGGAGCAGCTGTG	3600
OY	3602	GGCTCAGTTCGGGGGGAAGCATCAACAAAGGCAATCCCAAGCAACAAGGAGTCCCTCGGAC	3661
Db	3601	GGCTCAGTTCGGGGGGAAGCATCAACAAAGGCAATCCCAAGCAACAAGGAGTCCCTCGGAC	3660
OY	3662	AGCGCCATCACAATACCGGGGCTTCATCAACCAACGAGCAGCGCAAGTGAAGTCTGTGAAG	3721
Db	3661	AGCGCCATCACAATACCGGGGCTTCATCAACCAACGAGCAGCGCAAGTGAAGTCTGTGAAG	3720
OY	3722	GGCAACATCACACGATCATATGGCGAGAGACAGCCGAGTGGCTTGGAACCGGGCGGGAG	3781
Db	3721	GGCAACATCACACGATCATATGGCGAGAGACAGCCGAGTGGCTTGGAACCGGGCGGGAG	3780
OY	3782	GACAGCCTGCCAAGGCGCAGTCACTTAACGAAGGCAAGAAAGGCGCAAGTCTTGCTCAT	3841
Db	3781	GACAGCCTGCCAAGGCGCAGTCACTTAACGAAGGCAAGAAAGGCGCAAGTCTTGCTCAT	3840
OY	3842	GAGGGTGGCATGTCTGTGACCCAGTGTCCAAAGAGGACGGCAGAAAGAGCTCAGAGACC	3901

Db	3641	GAGGATGGCAATGTCTGTGACCCCACTGCTCCAGAGAGACGGCAGAAAGACTCAGAACCC	39000
Qy	3902	CCCCATAGACGCGCGCCCCCAAGCCGACCTATGACATGATGAGAGGCGCGGTGGGACGA	3961
Db	3901	CCCCATGAGACGCGCGCCCCCAAGCCGACCTATGACATGATGAGAGGCGCGGTGGGACGA	39600
Qy	3962	GCCATCTTCCACGCGACACATCGAAGGTCATCGAGGCCGTGCGCATCCCGCCGACCGACAC	40221
Db	3961	GCCATCTTCCACGCGACACATCGAAGGTCATCGAGGCCGTGCGCATCCCGCCGACCGACAC	40200
Qy	4022	AGCCCCACCACTCCAAAGACACACCACTCCGCGGGTCCATCAACAAGGATCCCT	4081
Db	4021	AGCCCCACCACTCCAAAGACACACCACTCCGCGGGTCCATCAACAAGGATCCCT	40800
Qy	4082	CGGTCCTACGTGAGAGCAACAGAGACCTACCTGCGTGGAGAGGCCAAAGCTCTAAAGCGG	41411
Db	4081	CGGTCCTACGTGAGAGCAACAGAGACCTACCTGCGTGGAGAGGCCAAAGCTCTAAAGCGG	41400
Qy	4142	GAGGGCACGCTCTCGCCCCCACCAGCTTACGGGACCTGACCGAGGCTTACAAAGCCAG	4201
Db	4141	GAGGGCACGCTCTCGCCCCCACCAGCTTACGGGACCTGACCGAGGCTTACAAAGCCAG	42000
Qy	4202	GCCCTGGGCCCCCTGAAAGCTGAAGCCGAGCCCATAGGGGCTGGAGGACAGGTAAAGGAG	4261
Db	4201	GCCCTGGGCCCCCTGAAAGCTGAAGCCGAGCCCATAGGGGCTGGAGGACAGGTAAAGGAG	4260
Qy	4262	GCGGCGCGCTCCATCATATGAGATCCCGCGGAGGCTGGCGGACACGCGCCGAGCTGCC	43211
Db	4261	GCGGCGCGCTCCATCATATGAGATCCCGCGGAGGCTGGCGGACACGCGCCGAGCTGCC	43200
Qy	4322	CTGGCCCCCGCGCGCTCCAGAGAGGCTCCATCAACGAGGACCCCGCTCAAGTACGAC	43811
Db	4321	CTGGCCCCCGCGCGCTCCAGAGAGGCTCCATCAACGAGGACCCCGCTCAAGTACGAC	43800
Qy	4382	ACCGGGCGGTCCACACATGCGCTCCAAAAGACGACGATCGCTCCCTCATGCGAGGCC	44411
Db	4381	ACCGGGCGGTCCACACATGCGCTCCAAAAGACGACGATCGCTCCCTCATGCGAGGCC	44400
Qy	4442	GCGCGGAGCTGTCCACCCGCTGACCCCGCTGAGTGTGATGCGCGACGCGCCGCGGACCTGGA	4501
Db	4441	GCGCGGAGCTGTCCACCCGCTGACCCCGCTGAGTGTGATGCGCGACGCGCCGCGGACCTGGA	45000
Qy	4502	CGTGCCTGTACGAGAGAGGCTTGAAGACCGCGCACGAGACCGCACGCTCGGGGGGCG	4561
Db	4501	CGTGCCTGTACGAGAGAGGCTTGAAGACCGCGCACGAGACCGCACGCTCGGGGGGCG	45600
Qy	4562	TTCATTGCGCGCGCGCGCGCGGTATTGTGCTGAGCTGGGTAGCCGCGGACAAAGCCCC	46221
Db	4561	TTCATTGCGCGCGCGCGCGCGGTATTGTGCTGAGCTGGGTAGCCGCGGACAAAGCCCC	46200
Qy	4622	CTGACCTATGAGAGACACGCGGACACCTTTGCGCGGACCTCCACGAGGTTGCGCCGCTG	4681
Db	4621	CTGACCTATGAGAGACACGCGGACACCTTTGCGCGGACCTCCACGAGGTTGCGCCGCTG	46800
Qy	4682	ACCATGCGGAGGACCAACGCGCGCTTACAGAGAGGACGCTTTGTCAGAGGACGATCC	47411
Db	4681	ACCATGCGGAGGACCAACGCGCGCTTACAGAGAGGACGCTTTGTCAGAGGACGATCC	47400
Qy	4742	CAGGACCGAAAGCTGACGTGAGCGCTTCGTGAGATCGCAAATGCCGACACGACCGCTG	4801
Db	4741	CAGGACCGAAAGCTGACGTGAGCGCTTCGTGAGATCGCAAATGCCGACACGACCGCTG	48000
Qy	4802	CCCGAGACCAACCAACCCCATCTGCGCCCTAGAGACCTGCTTGGGGGGGTGATGAGCG	4861
Db	4801	CCCGAGACCAACCAACCCCATCTGCGCCCTAGAGACCTGCTTGGGGGGGTGATGAGCG	48600
Qy	4862	GTCGACCTGATATGACACACATCCCTCGGCTTTCGACCCCATCTCATATCCCGCGGCG	4921
Db	4861	GTCGACCTGATATGACACACATCCCTCGGCTTTCGACCCCATCTCATATCCCGCGGCG	49200
Qy	4922	ATCCCTTGTGACGCGAGCGCTGCTTACTCTGCCCCGACCTTGGCCCCCAACCCACCC	4981



Db 4921 ATCCCTCTGAGCAGCCGCTGCTTACTACCTGCCCCGACACTGGCCCCAAACCCACC 4980  
Qy 4982 TACCCGCACTGTATCCCACTTACTCATCTCCGGGCTAACCCGACAGCGGGGCTGGAG 5041  
Db 4981 TACCCGCACTGTATCCCACTTACTCATCTCCGGGCTAACCCGACAGCGGGGCTGGAG 5040  
Qy 5042 AACCCGCACTGTATCCCACTTACTCATCTCCGGGCTAACCCGACAGCGGGGCTGGAG 5101  
Db 5041 AACCCGCACTGTATCCCACTTACTCATCTCCGGGCTAACCCGACAGCGGGGCTGGAG 5100  
Qy 5102 ACCGCACTGTATCCCACTTACTCATCTCCGGGCTAACCCGACAGCGGGGCTGGAG 5161  
Db 5101 ACCGCACTGTATCCCACTTACTCATCTCCGGGCTAACCCGACAGCGGGGCTGGAG 5160  
Qy 5162 GCACTCACTACGCTGCGGGTCCCGAGGATCATGACCTGCTCCCAAGTGCACACTG 5221  
Db 5161 GCACTCACTACGCTGCGGGTCCCGAGGATCATGACCTGCTCCCAAGTGCACACTG 5220  
Qy 5222 CCTGTGCTGTGCCCCGACACCAAGGACCCGACCGGCACTGACCGGCTTGTCTAC 5281  
Db 5221 CCGTGTGCTGTGCCCCGACACCAAGGACCCGACCGGCACTGACCGGCTTGTCTAC 5280  
Qy 5282 CTCCCACTGCGGCCCCGACCGCTTACGAGCGCCGACAGAGCTTCCCACTTCCCAAGA 5341  
Db 5281 CTCCCACTGCGGCCCCGACCGCTTACGAGCGCCGACAGAGCTTCCCACTTCCCAAGA 5340  
Qy 5342 GGTCCCACTGCTTGAACAAACCAACCAAGTCTCTGCTCGAGGGGAGCGAGACCGG 5401  
Db 5341 GGTCCCACTGCTTGAACAAACCAACCAAGTCTCTGCTCGAGGGGAGCGAGACCGG 5400  
Qy 5402 GATCGAGACGGGACCGGGATCGGAGCGGAGAAAGTCCATCTTCACTTCAACGAGAG 5461  
Db 5401 GATCGAGACGGGACCGGGATCGGAGCGGAGAAAGTCCATCTTCACTTCAACGAGAG 5460  
Qy 5462 GTGAGACAGCACCCATCTTGAGACCTGTGTACAGACCAAGACGGGACAGCGGACAG 5521  
Db 5461 GTGAGACAGCACCCATCTTGAGACCTGTGTACAGACCAAGACGGGACAGCGGACAG 5520  
Qy 5522 AACGGGGGGGTGGGGGACAGACCGGCCCCGCTCCCACTCCCAATGCCACAGACAG 5581  
Db 5521 AACGGGGGGGTGGGGGACAGACCGGCCCCGCTCCCACTCCCAATGCCACAGACAG 5580  
Qy 5582 TCGCCCATCTCCCTGGAGACCGAGATGCTCTCAGACAGAGACCGAGTGTCTTCAAC 5641  
Db 5581 TCGCCCATCTCCCTGGAGACCGAGATGCTCTCAGACAGAGACCGAGTGTCTTCAAC 5640  
Qy 5642 ACAGGATGAAGGATATCATCAACCGGTGTGAGACCGACGACCGGCTTGAAGTCC 5701  
Db 5641 ACAGGATGAAGGATATCATCAACCGGTGTGAGACCGACGACCGGCTTGAAGTCC 5700  
Qy 5702 ACTTCGACCTCTCAACCGGTGTGAGACCGAGTGTCTTCAACCTTGCACCTTGCCTCA 5761  
Db 5701 ACTTCGACCTCTCAACCGGTGTGAGACCGAGTGTCTTCAACCTTGCACCTTGCCTCA 5760  
Qy 5762 CTGGGGGAGACCTGTGAGGGGTCTACCTTACCTCATGAGAGCGGCTTGTGAGGCAAG 5821  
Db 5761 CTGGGGGAGACCTGTGAGGGGTCTACCTTACCTCATGAGAGCGGCTTGTGAGGCAAG 5820  
Qy 5822 GAGGCCCCCGGGGTGCGCCGGCGAGAGCGGCCCCGAGAGACACCGGCACTGCTTCTC 5881  
Db 5821 GAGGCCCCCGGGGTGCGCCGGCGAGAGCGGCCCCGAGAGACACCGGCACTGCTTCTC 5880  
Qy 5882 GGCMAAGCCCCAGCCCGCTTCGGGGCTGAGACCGGCTCTCTCCCGCAGCAAGGGCTGAG 5941  
Db 5881 GGCMAAGCCCCAGCCCGCTTCGGGGCTGAGACCGGCTCTCTCCCGCAGCAAGGGCTGAG 5940  
Qy 5942 CCGCGGCCCCCTAGTGTCTCTGTCTCTGAGCAAGGACCATGCGCGCACCCCTTGAGAG 6001  
Db 5941 CCGCGGCCCCCTAGTGTCTCTGTCTCTGAGCAAGGACCATGCGCGCACCCCTTGAGAG 6000  
Qy 6002 AACCTTGACCTTCAACGAGCGACCGGAGCCCGCGGCGCACTTGTGAGGCTTGAGAC 6061  
Db 6001 AACCTTGACCTTCAACGAGCGACCGGAGCCCGCGGCGCACTTGTGAGGCTTGAGAC 6060

Qy 6062 CCGCACCGGGAAAAGACTCAAAAGTAAACCTTTTCATCCAGAACTGAACTCGGTCT 6121  
Db 6061 CCGCACCGGGAAAAGACTCAAAAGTAAACCTTTTCATCCAGAACTGAACTCGGTCT 6120  
Qy 6122 CTGGGTTACAGGCGAGGCTTACAGCCCCGAAAGGGGTGAGGCTTCAAGCTTGTGAGC 6181  
Db 6121 CTGGGTTACAGGCGAGGCTTACAGCCCCGAAAGGGGTGAGGCTTCAAGCTTGTGAGC 6180  
Qy 6182 TCACCCAGCTGACCCCAAGCAAGGGGGTCCCAAGCACTGTGAGAGGCTGCAAGAGC 6241  
Db 6181 TCACCCAGCTGACCCCAAGCAAGGGGGTCCCAAGCACTGTGAGAGGCTGCAAGAGC 6240  
Qy 6242 CACTTGAAGGGGAGGTGCGGCTCAAGACGAGGCCCCGTGAAGCTTGGCGGGAGGCT 6301  
Db 6241 CACTTGAAGGGGAGGTGCGGCTCAAGACGAGGCCCCGTGAAGCTTGGCGGGAGGCT 6300  
Qy 6302 GCCCACTTCCCACTGCGGCTGCTGTGAGAGCAAGCCTTCTTCAAGCTTGTGCTC 6361  
Db 6301 GCCCACTTCCCACTGCGGCTGCTGTGAGAGCAAGCCTTCTTCAAGCTTGTGCTC 6360  
Qy 6362 CAGACCGCCCCAGGGGTCAAGGTACAGCGGGGTGTACCTGTGCGCAGACATCAAT 6421  
Db 6361 CAGACCGCCCCAGGGGTCAAGGTACAGCGGGGTGTACCTGTGCGCAGACATCAAT 6420  
Qy 6422 GAGGTATACACAGGACTACACCGGCGACACCCACAGAGCTCAGCGCACCCCTGCCC 6481  
Db 6421 GAGGTATACACAGGACTACACCGGCGACACCCACAGAGCTCAGCGCACCCCTGCCC 6480  
Qy 6482 GCCCCCTTACTCTTCTCCGTGGGGCAGCTGCCCCGTCTGTGACCTTCCGCGGCCACC 6541  
Db 6481 GCCCCCTTACTCTTCTCCGTGGGGCAGCTGCCCCGTCTGTGACCTTCCGCGGCCACC 6540  
Qy 6542 AGTGAACCTTACTCTCCGCCCCCGACATGTGTGCCCCGGGCTTCCGCCCCCAGC 6601  
Db 6541 AGTGAACCTTACTCTCCGCCCCCGACATGTGTGCCCCGGGCTTCCGCCCCCAGC 6600  
Qy 6602 GAAAGGGGCAAGAGGTCTCAGAGCCAAACAGTCCGCTTGTGGGTGTGAGTGAAGC 6661  
Db 6601 GAAAGGGGCAAGAGGTCTCAGAGCCAAACAGTCCGCTTGTGGGTGTGAGTGAAGC 6660  
Qy 6662 GGTATTTGAACCTGTGTCCCAACCGAGGGCATGACGAGCCAGAGGACTTCCGGAAGTCT 6721  
Db 6661 GGTATTTGAACCTGTGTCCCAACCGAGGGCATGACGAGCCAGAGGACTTCCGGAAGTCT 6720  
Qy 6722 GTGTACCCGCTGTGTACCGGGATGGGGAAACAGACGAGCCCAAGATGGGCTTCAAG 6781  
Db 6721 GTGTACCCGCTGTGTACCGGGATGGGGAAACAGACGAGCCCAAGATGGGCTTCAAG 6780  
Qy 6782 TCTCCAGGCAACACGAGCCGACGCGCTTCTTCAAGAGCTGACGAGAGCACTCC 6841  
Db 6781 TCTCCAGGCAACACGAGCCGACGCGCTTCTTCAAGAGCTGACGAGAGCACTCC 6840  
Qy 6842 GCCATGTCAATTCAGAGAGAGAGATCAACAGAGCTGAACCCAGACCCGGAAT 6901  
Db 6841 GCCATGTCAATTCAGAGAGAGAGATCAACAGAGCTGAACCCAGACCCGGAAT 6900  
Qy 6902 GAGCTGAATCAATATCAGCGAGCTGGGACGAGATTTCAATATGCCCCCATCAC 6961  
Db 6901 GAGCTGAATCAATATCAGCGAGCTGGGACGAGATTTCAATATGCCCCCATCAC 6960  
Qy 6962 GGAACAGGCTTATGACCTATGAGAGCGAGGGTGCAGAGAACTATGCCAGACCAATG 7021  
Db 6961 GGAACAGGCTTATGACCTATGAGAGCGAGGGTGCAGAGAACTATGCCAGACCAATG 7020  
Qy 7022 GGGCTGAGAGGCAATATTGAAGGCACTCATGGGTAAATATGACCAAGTGGAGAGTCC 7081  
Db 7021 GGGCTGAGAGGCAATATTGAAGGCACTCATGGGTAAATATGACCAAGTGGAGAGTCC 7080  
Qy 7082 CCGCGGCTCAGGCGCAATGCTTTTAACTTGAATGCAAGTGCAGGCTGCGCGTGCCT 7141  
Db 7081 CCGCGGCTCAGGCGCAATGCTTTTAACTTGAATGCAAGTGCAGGCTGCGCGTGCCT 7140



QY 7142 ATGCCATTAACCGCTGCTGACGAGACGAGTGAACAACACTCACTCCGACAGTGGCGGC 7201  
DB 7141 ATGCCATTAACCGCTGCTGACGAGACGAGTGAACAACACTCACTCCGACAGTGGCGGC 7200  
QY 7202 GGGAAAGGCGAAGGCTCTGGGAGAGACCGACGAGCGAAAAAGCCAAAGTCCCGGGCCCGGGC 7261  
DB 7201 GGGAAAGGCGAAGGCTCTGGGAGAGACCGACGAGCGAAAAAGCCAAAGTCCCGGGCCCGGGC 7260  
QY 7262 CTGGCATCTGGGAGACCGGCAACCTCTGTCTCTCTCACTGAGTCACTCGAGGAGAACTGCAAC 7321  
DB 7261 CTGGCATCTGGGAGACCGGCAACCTCTGTCTCTCTCACTGAGTCACTCGAGGAGAACTGCAAC 7320  
QY 7322 CGCCGAGAGCGCGCTCAACCAACCGGCTGTGGAGAGAACAGGCTCTCGTCCGAGATTCACG 7381  
DB 7321 CGCCGAGAGCGCGCTCAACCAACCGGCTGTGGAGAGAACAGGCTCTCGTCCGAGATTCACG 7380  
QY 7382 CCATTCCCTTCAACACCCCTGATCATGAGGAGTGAAGGAGGAGTGTAGTATGGGCTTCCCAACC 7441  
DB 7381 CCATTCCCTTCAACACCCCTGATCATGAGGAGTGAAGGAGGAGTGTAGTATGGGCTTCCCAACC 7440  
QY 7442 CCACCGGAGCTCCCGCGGAGCGGAGCCCTCGCTGAGCCCCACACGCTGAGAGCAG 7501  
DB 7441 CCACCGGAGCTCCCGCGGAGCGGAGCCCTCGCTGAGCCCCACACGCTGAGAGCAG 7500  
QY 7502 GAGCCCAAGCCACTGCTCTGCTGCAATGAGAGCACTTCCGACAGGAGTGA 7555  
DB 7501 GAGCCCAAGCCACTGCTCTGCTGCAATGAGAGCACTTCCGACAGGAGTGA 7554

## RESULT 9

US-09-819-104A-3  
; Sequence 3, Application US/09819104A  
; Publication No. US20030027137A1  
; GENERAL INFORMATION:  
; APPLICANT: Chem. J. Don  
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES  
; FILE REFERENCE: UMG-030  
; CURRENT APPLICATION NUMBER: US/09/819,104A  
; CURRENT FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: 60/193,138  
; PRIOR FILING DATE: 2000-03-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 7521  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(7521)  
US-09-819-104A-3

ABN

Query Match 85.6%; Score 7334.6; DB 3; Length 7521;  
Best Local Similarity 98.5%; Pred. No. 0;  
Matches 7463; Conservative 0; Mismatches 34; Indels 78; Gaps 3;

QY 2 ATGTGGGCTCCACACAGCTGTGGACACAGAGTGAAGGAGCCACTGAGCCCGTACCG 61  
DB 1 ATGTGGGCTCCACACAGCTGTGGACACAGAGTGAAGGAGCCACTGAGCCCGTACCG 60  
QY 62 CCCACAGCTTTCTTACCAGTGAAGATGCGCCGAGCGACACGAGCGTGGGCTCTCTG 121  
DB 61 CCCACAGCTTTCTTACCAGTGAAGATGCGCCGAGCGACACGAGCGTGGGCTCTCTG 120  
QY 122 GAGTACAGACCACTCCCGGCACTATGCTCTCCACCTGTGCGCGGGCTTCATTCAG 181  
DB 121 GAGTACAGACCACTCCCGGCACTATGCTCTCCACCTGTGCGCGGGCTTCATTCAG 180  
QY 182 CCCAGCGGCGAGGCGCTCCCTGCTGTGAGTTCCAGCCGAGGAATGAAGTCCCG 241  
DB 181 CCCAGCGGCGAGGCGCTCCCTGCTGTGAGTTCCAGCCGAGGAATGAAGTCCCG 240

QY 242 GAGCTTCACTTGGCGCCAGAGATCCCATCATATCTGCGCCGAGCTGGGAGATCAGAGATG 301  
DB 241 GAGCTTCACTTGGCGCCAGAGATCCCATCATATCTGCGCCGAGCTGGGAGATCAGAGATG 300  
QY 302 GAGTTCAATTGAAGCAAGCGCTTCGCTAGAGCTGCTGCAACCCCTGCTGGAGCG 361  
DB 301 GAGTTCAATTGAAGCAAGCGCTTCGCTAGAGCTGCTGCAACCCCTGCTGGAGCG 360  
QY 362 TCACCCCTGCTGGCCACGAGCCAGCTTGGAGATCTGAAGACTTCACCAAGACCGTAGC 421  
DB 361 TCACCCCTGCTGGCCACGAGCCAGCTTGGAGATCTGAAGACTTCACCAAGACCGTAGC 420  
QY 422 CTGACGGGCAAGCTTGAACCCGCTGTCTCCCGCAGCGCCCCCGCAGACTGAGCTG 481  
DB 421 CTGACGGGCAAGCTTGAACCCGCTGTCTCCCGCAGCGCCCCCGCAGACTGAGCTG 480  
QY 482 GAGCTGTGGCGCCACGAGCTGTCCAAGGAGAGCTGATCCAGAACATGAGACCGCTGAC 541  
DB 481 GAGCTGTGGCGCCACGAGCTGTCCAAGGAGAGCTGATCCAGAACATGAGACCGCTGAC 540  
QY 542 CGAGAGATCAACATGTGTAGAGCAAGAGATCTTAAGCTGAAGAAAGACAGCAAGCTG 601  
DB 541 CGAGAGATCAACATGTGTAGAGCAAGAGATCTTAAGCTGAAGAAAGACAGCAAGCTG 600  
QY 602 GAGGAGAGGCTTGGCCAAAGCCCGCCGAGCTGAGAAAGCCCGTGTCAACGCGCCCATGAG 661  
DB 601 GAGGAGAGGCTTGGCCAAAGCCCGCCGAGCTGAGAAAGCCCGTGTCAACGCGCCCATGAG 660  
QY 662 TCGAAGCACCGGAGCTGTGAGATCATCTACGACGAGAACCGGAGAAAGGCTGAGCT 721  
DB 661 TCGAAGCACCGGAGCTGTGAGATCATCTACGACGAGAACCGGAGAAAGGCTGAGCT 720  
QY 722 GCACATCGAATTTTGAAGGCTGTGGGCCCCGAGGTGAGCTGCGCTGTACCAAGGCC 781  
DB 721 GCACATCGAATTTTGAAGGCTGTGGGCCCCGAGGTGAGCTGCGCTGTACCAAGGCC 780  
QY 782 TCCGACACCGGAGATCATGAGAAACATCAAAATTAACAGGCGATCGGAGAAAGCTA 841  
DB 781 TCCGACACCGGAGATCATGAGAAACATCAAAATTAACAGGCGATCGGAGAAAGCTA 840  
QY 842 ATCTTGTACTTCAAGAGGAGATCAAGCTGCGGAAACAATGGAAGCAGAGTTCTGCAG 901  
DB 841 ATCTTGTACTTCAAGAGGAGATCAAGCTGCGGAAACAATGGAAGCAGAGTTCTGCAG 900  
QY 902 CGCTATGACCAAGCTCATGAGAGCCCTTGGAAAAAAGGTGAGCGCATGAAAAACAACCG 961  
DB 901 CGCTATGACCAAGCTCATGAGAGCCCTTGGAAAAAAGGTGAGCGCATGAAAAACAACCG 960  
QY 962 CGCGGCGGGCCAGGAGAGCAAGGTGCGGAGTCTACGAAAGCAAGTTCTCTGAGATC 1021  
DB 961 CGCGGCGGGCCAGGAGAGCAAGGTGCGGAGTCTACGAAAGCAAGTTCTCTGAGATC 1020  
QY 1022 CGCAAGCAGCGGAGCTGACGAGAGCGCATGACAGAGCAGAGTGGGCGACGCGGAGTGG 1081  
DB 1021 CGCAAGCAGCGGAGCTGACGAGAGCGCATGACAGAGCAGAGTGGGCGACGCGGAGTGG 1077  
QY 1082 CTGTCCATGTGGCGCCCGCCGAGCGAGCAGAGTGTGAGAGATCATGAGCTCTCA 1141  
DB 1078 CTGTCCATGTGGCGCCCGCCGAGCGAGCAGAGTGTGAGAGATCATGAGCTCTCA 1137  
QY 1142 GAGCAGAGAAACTTGGAGAAAGATGCGCCAGCTGCGCGCTGATCCCGCCCATGCTGAC 1201  
DB 1138 GAGCAGAGAAACTTGGAGAAAGATGCGCCAGCTGCGCGCTGATCCCGCCCATGCTGAC 1197  
QY 1202 GAGCTGACCAAGCGGATCAAGTTTCAACATGAACGAGGCTTATGAGCGGACCCCATG 1261  
DB 1198 GAGCTGACCAAGCGGATCAAGTTTCAACATGAACGAGGCTTATGAGCGGACCCCATG 1257  
QY 1262 AAGGTGTACAAAGACCGCAGGTCATGAACATGTGAGTGAAGCAGAGAGAGACCTTC 1321  
DB 1258 AAGGTGTACAAAGACCGCAGGTCATGAACATGTGAGTGAAGCAGAGAGAGACCTTC 1317  
QY 1322 CGGGAAGATTCATGACAGATCCCAAGAACTTTGGCTGATCGCATCTCTGAGAGAG 1381

```
Db      1318 CGGAGAAAGTTCAATGACAGATCCCAAGAACTTGGCCATGATCCATCATCTCTGAGAGG 1377
Qy      1382 AAGACAGTGCCTAAGTGGCCCTCTTATTAATTAATCTAATAAGATGAATGAATCTAAG 1441
Db      1378 AAGACAGTGCCTAAGTGGCCCTCTTATTAATTAATCTAATAAGATGAATGAATCTAAG 1437
Qy      1442 AGCCTGATGAGACGAGACTTCCGCGCGCGGCAAGAGCCAGACGAAACAACAGCAGCAG 1501
Db      1438 AGCCTGATGAGACGAGACTTCCGCGCGCGGCAAGAGCCAGACGAAACAACAGCAGCAG 1497
Qy      1502 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1561
Db      1498 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1557
Qy      1562 GATGAGAAAGAGAAAGAAAGAGAGCGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621
Db      1558 GATGAGAAAGAGAAAGAAAGAGAGCGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1617
Qy      1622 GACAAAGAAAGAACTCTCTCAAGAGAAAGACAGACGACCTCAGGGAGAGACAAAGCAGAG 1681
Db      1618 GACAAAGAAAGAACTCTCTCAAGAGAAAGACAGACGACCTCAGGGAGAGACAAAGCAGAG 1677
Qy      1682 AAGAGAGCTGTGGCTTCCAAAGGCGCGCAAACTGCCAACGCGAGAGAGACGCAAAAGCG 1741
Db      1678 AAGAGAGCTGTGGCTTCCAAAGGCGCGCAAACTGCCAACGCGAGAGAGACGCAAAAGCG 1737
Qy      1742 CGCATCAACCGCTCAATTGGCTAATGAGGCCAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1801
Db      1738 CGCATCAACCGCTCAATTGGCTAATGAGGCCAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1797
Qy      1802 AGCGCGAGCTGGCTCCATGAGCTGAATGAGATTTCTGCTGAGACAGAAAGAAAGT 1861
Db      1798 AGCGCGAGCTGGCTCCATGAGCTGAATGAGATTTCTGCTGAGACAGAAAGAAAGT 1857
Qy      1862 GAAACAGCCAAAGAAAGATCTCTTGAACAACGCGCGCAACTGTGCGCATCGCCGAGATG 1921
Db      1858 GAAACAGCCAAAGAAAGATCTCTTGAACAACGCGCGCAACTGTGCGCATCGCCGAGATG 1917
Qy      1922 GTGGGCTCCAAAGCTGTGTGCAAGTGAAGACTTTCTTCAACTTCAAGAAAGAGCAG 1981
Db      1918 GTGGGCTCCAAAGCTGTGTGCAAGTGAAGACTTTCTTCAACTTCAAGAAAGAGCAG 1977
Qy      1982 AACCTGATGAGATCTTTGACGAGCACAAGCTGAATGAGAGAGAGAGAGAGAGAGAGAGAG 2041
Db      1978 AACCTGATGAGATCTTTGACGAGCACAAGCTGAATGAGAGAGAGAGAGAGAGAGAGAGAG 2037
Qy      2042 AGGAAAGAGAAAGAAAGCGCGCGCGCGCGCGAGAGAGAGCTCATTTCCGCGCCGTGTG 2101
Db      2038 AGGAAAGAGAAAGAAAGCGCGCGCGCGCGCGAGAGAGAGCTCATTTCCGCGCCGTGTG 2097
Qy      2102 GAGGATGAGAGATGAGAGCGCTGCGCGGTGAGCGGAAATGAGAGAGAGATGTGTGAGAG 2161
Db      2098 GAGGATGAGAGATGAGAGCGCTGCGCGGTGAGCGGAAATGAGAGAGAGATGTGTGAGAG 2157
Qy      2162 GCTGAAGCCTTACATCCTCTGAGGAATGAGGTCCCAAGAGGGAAATGAGTGGCCCAAGCC 2221
Db      2158 GCTGA-----AGCC 2166
Qy      2222 ACTGTCAACAACAGCTCAGACACGAGAGATCCCTCTCTCACTGAGGCGGCAAG 2281
Db      2167 ACTGTCAACAACAGCTCAGACACGAGAGATCCCTCTCTCACTGAGGCGGCAAG 2226
Qy      2282 GACACAGGCGAAGATGGGCCCAAGCCCCCAGCCACCTTGGGCGCGCAGCGGCGCACCCCA 2341
Db      2227 GACACAGGCGAAGATGGGCCCAAGCCCCCAGCCACCTTGGGCGCGCAGCGGCGCACCCCA 2286
Qy      2342 GGGCCCAACCAACCAACGAGAGACATCCCGGGCCCCCAATGAGCCCAACCCGCGCTCT 2401
Db      2287 GGGCCCAACCAACCAACGAGAGACATCCCGGGCCCCCACTAGTGTCAACCCGCGCTCT 2346
Qy      2402 GAAGCCACGAGAGCCCTTACGCCCCCAACGAGACCCCAATCGCCCTGTGCACTCTCTCT 2461
Db      2347 GAAGCCACCTTACGCCCTTACGCCCCCAACGAGACCCCAATTTCCCTTCACTCTCTCT 2406
Qy      2462 GTGGTCCCCAAGAGAGAGAGAGAGAGACCGCAGACAGCGCCCCCAAGTGAAGAGGG 2521
Db      2407 GTGGTCCCCAAGAGAGAGAGAGAGAGACCGCAGACAGCGCCCCCAAGTGAAGAGGG 2466
Qy      2522 GAGAGACAGAGAGCCCCCGCGCTGAGAGCTTGGCAATGAGACAAGAGAGAGCGAGAG 2581
Db      2467 GAGAGACAGAGAGCCCCCGCGCTGAGAGCTTGGCAATGAGACAAGAGAGAGCGAGAG 2526
Qy      2582 CCCGTCAAGAGCGATGTCACGAGAGAGACCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2641
Db      2527 CCCGTCAAGAGCGATGTCACGAGAGAGACCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2586
Qy      2642 GAGGCCGTGAGAGCCAGCGCGCAGAGGGGGCTCAAGGCAAGAAAGAGAGAGAGAGAGAG 2701
Db      2587 GAGGCCGTGAGAGCCAGCGCGCAGAGGGGGCTTCAAGGCAAGAAAGAGAGAGAGAGAGAG 2646
Qy      2702 GGCAGGGGCCACACTTGCACAAAGACTCGAGCGCCCCCAGAGACAGCACTTCAATGTCTAC 2761
Db      2647 GGCAGGGGCCACCAACAGCAAGAGCTCGAGCGCCCCCAGAGACAGCACTTCAATGTCTAC 2706
Qy      2762 TGCATGTGACAGAGAGGTGATGAGGCTCGAGGCGCGGACAGAAACCGGCTGTCTCCCA 2821
Db      2707 TGCATGTGACAGAGAGGTGATGAGGCGCGGCGGACAGAAACCGGCTGTCTCCCA 2766
Qy      2822 AGGCCAGGCTCTCTCAACCCGACTGGCGACCCCGGGCCCAATGCTCACCCAGAAAGCA 2881
Db      2767 AGGCCAGGCTCTCTCAACCCGACTGGCGACCCCGGGCCCAATGCTCTCACCCAGAAAGCA 2826
Qy      2882 CTGGAACCTGAAGAGCTGAAAGACGAGAGCGGCTGCATCCCCCCTCAAGATCAACAA 2941
Db      2827 CTGGAACCTGAAGAGCTGAAAGACGAGAGCGGCTGCATCCCCCCTCAAGATCAACAA 2886
Qy      2942 GTTCATGAGCCCCCGGGAGAGACGAGCTCCCAACAGCAGCTTCCCAAGCCCAACCG 3001
Db      2887 GTTCATGAGCCCCCGGGAGAGACGAGCTCCCAACAGCAGCTTCCCAAGCCCAACCG 2946
Qy      3002 CCAACGCCAAACCTTGAGCGCGAGAGCGAGCGCCCTCAACAGCTTGGAGAGAGAGAGAGAG 3061
Db      2947 CCAACGCCAAACCTTGAGCGCGAGAGCGAGCGCCCTCAACAGCTTGGAGAGAGAGAGAGAG 3006
Qy      3062 GGCAGAGAGAGAGAGCGCGCACCCCGCGCAGAAAG----- 3100
Db      3007 GGCAGAGAGAGAGAGCGCGCACCCCGCGCAGAAAGCTGTGTCTTCC 3066
Qy      3101 ---GCCTTGCAGAGCGAGGCCCAAGAGCTGCTGGAGACCCCTTGTGAGACTTCCGGC 3157
Db      3067 CCAAGCTTGCAGAGCGAGGCCCAAGAGCTGCTGGAGAGACCCCTTGTGAGACTTCCGGC 3126
Qy      3158 CTGCCCTTCCCGTGGCCCCCGGTGAGTGAATCAAGGCTTCCCGGATGCCCGGAGCCCC 3217
Db      3127 CTGCCCTTCCCGTGGCCCCCGGTGAGTGAATCAAGGCTTCCCGGATGCCCGGAGCCCC 3186
Qy      3218 TCAGCCTTCTCTTACGCTCACAACCTGATCAACCACTGCCCCCTGGAGCTTCATGACTGCG 3277
Db      3187 TCAGCCTTCTCTTACGCTCACAACCTGATCAACCACTGCCCCCTGGAGCTTCATGACTGCG 3246
Qy      3278 CGGCGCGTCTGTGCGCGGCCACCCACATCTCCAAACCGGCTTCCCTCATCTCTGTGCC 3337
Db      3247 CGGCGCGTCTGTGCGCGGCCACCCACATCTCCAAACCGGCTTCCCTCATCTCTGTGCC 3306
Qy      3338 AAGCAACCCAGAGGTCTTCAAGAGGCAAAATGAGTCCATCTCCCAAGAAATGTGCTCCAG 3397
Db      3307 AAGCAACCCAGAGGTCTTCAAGAGGCAAAATGAGTCCATCTCCCAAGAAATGTGCTCCAG 3366
Qy      3398 CTCACGTCCCGTACTCAGAGATGCAAGAGCCCGGATGGGCGCTGTCAACATGGGGCTG 3457
Db      3367 CTCACGTCCCGTACTCAGAGATGCAAGAGCCCGGATGGGCGCTGTCAACATGGGGCTG 3426
Qy      3458 CCCCTGCCATGAGACCCCAAAAGCTGTGAGCCCTTCAAGGAGTGAAGAGAGAGAGCTG 3517
Db      3427 CCCCTGCCATGAGACCCCAAAAGCTGTGAGCCCTTCAAGGAGTGAAGAGAGAGAGCTG 3486
```

QY 3518 TCCCAACGGGGCTGAGCTGGCCACCGAGAGCTGGGGTGGCCACAGCCACGAGGCG 3577  
| | | | |  
Db 3487 TCCCAACGGGGCTGAGCTGGGGCCACCGAGAGCTGGGGTGGCCACAGCCACGAGGCG 3546  
| | | | |  
QY 3578 TCCGCTGAGAGGGGACAGCTCGGGCTGAGTTCCGGGGGGAAGATCCCAAGGCATT 3637  
| | | | |  
Db 3547 TCCGCTGAGAGGGGACAGCTCTGGGCTGAGTTCCGGGGGGAAGATCCCAAGGCATT 3606  
| | | | |  
QY 3638 CCACAGACACGGGGTCCCTCGACAGCGCCATCATACATACCGGGCTCCATCACCCACGCG 3697  
| | | | |  
Db 3607 CCACAGACACGGGGTCCCTCGACAGCGCCATCATACATACCGGGCTCCATCACCCACGCG 3666  
| | | | |  
QY 3698 ACGCACTGACGTCTCTGTACAGGGCACCATCACAGGATCATGGGCGAGACAGCCCG 3757  
| | | | |  
Db 3667 ACGCACTGACGTCTCTGTACAGGGCACCATCACAGGATCATGGGCGAGACAGCCCG 3726  
| | | | |  
QY 3758 AGTCGCTTGAACGGGGCCGGGAGGACAGCGCTGGCCCAAGGGCACTCATCAAGAGCG 3817  
| | | | |  
Db 3727 AGTCGCTTGAACGGGGCCGGGAGGACAGCGCTGGCCCAAGGGCACTCATCAAGAGCG 3786  
| | | | |  
QY 3818 AAGAGGGCCACGTCTTGTCTATGAGGGTGGCATGTCTGTACCCAGTCTCAAGAG 3877  
| | | | |  
Db 3787 AAGAGGGCCACGTCTTGTCTATGAGGGTGGCATGTCTGTACCCAGTCTCAAGAG 3846  
| | | | |  
QY 3878 ACGGCGAAGAGAGCTCAGGACCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGAC 3937  
| | | | |  
Db 3847 ACGGCGAAGAGAGCTCAGGACCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGAC 3906  
| | | | |  
QY 3938 ATGATGAGAGGGCGGGTGGGAGAGCATCTCTCAGCCGACGATGAGAGGTCTCATGGGC 3997  
| | | | |  
Db 3907 ATGATGAGAGGGCGGGTGGGAGAGCATCTCTCAGCCGACGATGAGAGGTCTCATGGGC 3966  
| | | | |  
QY 3998 CGTGCATCTCCGCCGAGAGGACACAGCCCCACACACTTCAAAAGACAGACCACTCCG 4057  
| | | | |  
Db 3967 CGTGCATCTCCGCCGAGAGGAGACAGCCCCACACACTTCAAAAGAGAGACCACTCCG 4026  
| | | | |  
QY 4058 GGGTCAATCACAAAGGATCCCTGGTCTTACGTGAGAGCACAGAGAGCTTACTTGGT 4117  
| | | | |  
Db 4027 GGGTCAATCACAAAGGATCCCTGGTCTTACGTGAGAGCACAGAGAGCTTACTTGGT 4086  
| | | | |  
QY 4118 CGGGAGGCGAAGCTCTAAAGGGGAGGGGCAAGCCTCCGCCCCCACCGGCTTCAAGGAG 4177  
| | | | |  
Db 4087 CGGGAGGCGAAGCTCTTAAAGGGGAGGGGCAAGCCTCCGCCCCCACCGGCTTCAAGGAG 4146  
| | | | |  
QY 4178 CTGACCGAGGCTTACAGACGCGAGCCCTGGGCCCCCTGAAGCTGAAGCCGCCATGAG 4237  
| | | | |  
Db 4147 CTGACCGAGGCTTACAGACGCGAGCCCTGGGCCCCCTGAAGCTGAAGCCGCCATGAG 4206  
| | | | |  
QY 4238 GGCCTGTGGCCACGATGAAGAGGCGGGCGGCTTCATCATGAGATCCCGCGAGAGAG 4297  
| | | | |  
Db 4207 GGCCTGTGTGGCCACGATGAAGAGGCGGGCGGCTTCATCATGAGATCCCGCGAGAGAG 4266  
| | | | |  
QY 4298 CTGCGGACACGCGCCGAGCTGGCCCCCTGGGCCCCGCTCAAGAGGGCTTCATCAG 4357  
| | | | |  
Db 4267 CTGCGGACACGCGCCGAGCTGGCCCCCTGGGCCCCGCTCAAGAGGGCTTCATCAG 4326  
| | | | |  
QY 4358 CAGGACACCCCGCTCAAGTACGACACCGGGCGGTCCACCATGTGCTCCAAAAAGACAGC 4417  
| | | | |  
Db 4327 CAGGACACCCCGCTCAAGTACGACACCGGGCGGTCCACCATGTGCTCCAAAAAGACAGC 4386  
| | | | |  
QY 4418 GTACGCTCTCTCATTCGAGAGCCCGGCGGAGCGTTCCACCCGTCACCCGCTGATGTG 4477  
| | | | |  
Db 4387 GTACGCTCTCTCATTCGAGAGCCCGGCGGAGCGTTCCACCCGTCACCCGCTGATGTG 4446  
| | | | |  
QY 4478 ATGGCGAGCGCCCGGGACCTGGAAGGTGCTGTCTACAGAGAGAGCTTGAAGCGGGCCA 4537  
| | | | |  
Db 4447 ATGGCGAGCGCCCGGGACCTGGAAGGTGCTGTCTACAGAGAGAGCTTGAAGCGGGCCA 4506  
| | | | |  
QY 4538 GGGACCGACAGAGCTCGGGGGGCTCATTTGCGCGGGCGCCCGGTCAATTTGTGCTGAG 4597  
| | | | |  
Db 4507 GGGACCGACAGAGCTCGGGGGGCTCATTTGCGCGGGCGCCCGGTCAATTTGTGCTGAG 4566  
| | | | |

QY 4558 CTGGGTAAAGCCCGGCAAGACCCCTTGAACCTATGAGACACAGGGGACACCTTTGCCGCG 4657  
| | | | |  
Db 4567 CTGGGTAAAGCCCGGCAAGACCCCTTGAACCTATGAGACACAGGGGACACCTTTGCCGCG 4626  
| | | | |  
QY 4658 CACCTCCCAAGAGTTCCGCCGTGACCATGCGGAGGCCACGCGCGCTGACGAGAGGGC 4717  
| | | | |  
Db 4627 CACCTCCCAAGAGTTCCGCCGTGACCATGCGGAGGCCACGCGCGCTGACGAGAGGGC 4686  
| | | | |  
QY 4718 AGCCTTTGCTCAGCAAGGATCCAGAGACGAAAGCTGACGTTCGACGCTCTGTGAGATC 4777  
| | | | |  
Db 4687 AGCCTTTGCTCAGCAAGGATCCAGAGACGAAAGCTGACGTTCGACGCTCTGTGAGATC 4746  
| | | | |  
QY 4778 GCGAAGTCCCGCACAGACCGGTGCCGAGACACCCACACCCCATCTCGCTTATGAG 4837  
| | | | |  
Db 4747 GCGAAGTCCCGCACAGACCGGTGCCGAGACACCCACACCCCATCTCGCTTATGAG 4806  
| | | | |  
QY 4838 CACCTGCTTGGGGGAGTGAAGTGGCGGTGACCTGTATGCAAGCACATCCCCCTGGCCTTC 4897  
| | | | |  
Db 4807 CACCTGCTTGGGGGAGTGAAGTGGCGGTGACCTGTATGCAAGCACATCCCCCTGGCCTTC 4866  
| | | | |  
QY 4898 GACCCCACTTCATACCCCGGAGCATCCCTGTGACGAGCGCTGCTACTTACTTGGCC 4957  
| | | | |  
Db 4867 GACCCCACTTCATACCCCGGAGCATCCCTGTGACGAGCGCTGCTACTTACTTGGCC 4926  
| | | | |  
QY 4958 GGAACCTGGCCCCCAACCCCATCTTACCCGACCTGTATCCACCTTACTTACTTGGCGC 5017  
| | | | |  
Db 4927 GGAACCTGGCCCCCAACCCCATCTTACCCGACCTGTATCCACCTTACTTACTTGGCGC 4986  
| | | | |  
QY 5018 TACCCCGACAGGGCGGCTGGAAGAACGGGACACATCATCATGATCATCATCTCG 5077  
| | | | |  
Db 4987 TACCCCGACAGGGCGGCTGGAAGAACGGGACACATCATCATGATCATCATCTCG 5046  
| | | | |  
QY 5078 CAGCAGATGCAACCAACAGGACACCGCATGCGCCAGCAGCTGATATGCTGAGGGGC 5137  
| | | | |  
Db 5047 CAGCAGATGCAACCAACAGGACACCGCATGCGCCAGCAGCTGATATGCTGAGGGGC 5106  
| | | | |  
QY 5138 CTCTGCCCCCGAGTCTCGCTGGACCTTCACTACGTGCGGGTCCCCGAGGACTATC 5197  
| | | | |  
Db 5107 CTCTGCCCCCGAGTCTCTGCTGGCACTTCACTACGTGCGGGTCCCCGAGGACTATC 5166  
| | | | |  
QY 5198 GACCTGTCCAAAGTGCACACTGTGCTGTGGTCCGGGCCCGGACACAGGGACCCCAAGC 5257  
| | | | |  
Db 5167 GACCTGTCCAAAGTGCACACTGTGCTGTGGTCCGGGCCCGGACACAGGGACCCCAAGC 5226  
| | | | |  
QY 5258 ACCGCTATGAGACCGGCTTGCCTTACCTCCCAACCGGCCCCAGCCCTTGAAGACCGCAC 5317  
| | | | |  
Db 5227 ACCGCTATGAGACCGGCTTGCCTTACCTCCCAACCGGCCCCAGCCCTTGAAGACCGCAC 5286  
| | | | |  
QY 5318 AGCAGCTCCCACTCTCCCAAGAGTCCAAACATTTGAAGAAACAAACCAACACAGTCC 5377  
| | | | |  
Db 5287 AGCAGCTCCCACTCTCCCAAGAGTCCAAACATTTGAAGAAACAAACCAACACAGTCC 5346  
| | | | |  
QY 5378 TCGTCCGAGCGGGAGCGGAGACCGGGATCGAAGGGGACCGGGATCCGGGAACGGGAAAG 5437  
| | | | |  
Db 5347 TCGTCCGAGCGGGAGCGGAGACCGGGATCGAAGGGGACCGGGATCCGGGAACGGGAAAG 5406  
| | | | |  
QY 5438 TCCATCTCTCAAGTCCACACAGACGGTGGAGCACCATCTGGAGACTTGTATACAGAG 5497  
| | | | |  
Db 5407 TCCATCTCTCAAGTCCACACAGACGGTGGAGCACCATCTGGAGACTTGTATACAGAG 5466  
| | | | |  
QY 5498 CAGAGCAGCGGACAGCGGAGCAGCGCGGGGGTGGGGGACAGACGCGCCCGCC 5557  
| | | | |  
Db 5467 CAGAGCAGCGGACAGCGGAGCAGCGCGGGGGTGGGGGACAGACGCGCCCGCC 5526  
| | | | |  
QY 5558 TCCCATCTCCCATGCGCCACACACATCCGCGCATCTCCCTCGGACCCAGAGGCTCTCAG 5617  
| | | | |  
Db 5527 TCCCATCTCCCATGCGCCACACACATCTGCGCATCTCCCTCGGACCCAGAGGCTCTCAG 5586  
| | | | |  
QY 5618 CAGAGACCCAGTGTCTTCAACAACAGGCAATGAAGGTATCATCACCGCTGTGAGGCC 5677  
| | | | |  
Db 5587 CAGAGACCCAGTGTCTTCAACAACAGGCAATGAAGGTATCATCACCGCTGTGAGGCC 5646  
| | | | |  
QY 5678 AGCAAGCCCAAGTCTGAGGTCAACTTCACTTCAACCCGTTGCGCCAGCTGCCA 5737  
| | | | |

|||||  
Db 5647 AGCAACCCACGCTCTGAGGTCCACCTCCACCCGCTGCGCCAGCTGCACGA 5706  
5738 TTCCGACCTGGCCACCCATGCGCCGACCTCGATGGGGCTTACCTTACCTC 5797  
5707 TTCCGACCTGGCCACCCATGCGCCGACCTCGATGGGGCTTACCTTACCTC 5766  
5798 ATGAGCCCGCTTGTGCTGCGCAAGAGGCCCGCGGGTCGCGCCGACAGCGGCCGA 5857  
5767 ATGAGCCCGCTTGTGCTGCGCAAGAGGCCCGCGGGTCGCGCCGACAGCGGCCGA 5826  
5858 GCAAGACCGGCGCATCTCTCTGCGCAAGGCCCGCGCTCCGCGCTGAGCGCCGC 5917  
5827 GCAAGACCGGCGCATCTCTCTGCGCAAGGCCCGCGCTCCGCGCTGAGCGCCGC 5886  
5918 TCCCTCCCGCAGCAAGGGCTCGAGGCCCGCGCTCTGTGCTCTTCTGCGCCAGCC 5977  
5887 TCCCTCCCGCAGCAAGGGCTCGAGGCCCGCGCTCTGTGCTCTTCTGCGCCAGCC 5946  
5978 ACCATGCGCCGCAACCTGCGAAGAACTGCGACTCACCGACGCGCCGCGCGCG 6037  
5947 ACCATGCGCCGCAACCTGCGAAGAACTGCGACTCACCGACGCGCCGCGCGCG 6006  
6038 GCGCCACCTGCTCGGCTCGGCTCGGACCGCGCAACCGGAAAAAGCTCAAGTAAACCTTTTCC 6097  
6007 GCGCCACCTGCTCGGCTCGGACCGCGCAACCGGAAAAAGCTCAAGTAAACCTTTTCC 6066  
6098 ATCCAGAACTGGAATCTCGGCTCTGCGGTACCAAGCGCAGCGCTACAGCCCCGAAAGG 6157  
6067 ATCCAGAACTGGAATCTCGGCTCTGCGGTACCAAGCGCAGCGCTACAGCCCCGAAAGG 6126  
6158 GTGAGACCGCTGAGCCCTGAGACTCAACCGCTGAGCCAGCAAGGGGCTCCCCAAG 6217  
6127 GTGAGACCGCTGAGCCCTGAGACTCAACCGCTGAGCCAGCAAGGGGCTCCCCAAG 6186  
6218 CACTGGAAGAGCTCGACAAAGAGCACTGAGAGGGGAGCTGCGGCCCAAGCAGCCAGG 6277  
6187 CACTGGAAGAGCTCGACAAAGAGCACTGAGAGGGGAGCTGCGGCCCAAGCAGCCAGG 6246  
6278 CCCGTGAACTTGGCGGGGAGCGGCCCACTCCCACTTGGGCGCGCTGTAGAGC 6337  
6247 CCCGTGAACTTGGCGGGGAGCGGCCCACTCCCACTTGGGCGCGCTGTAGAGC 6306  
6338 CAGCCCTGTCAGAGCCCGCTGCTCAGACCGGCCCAAGGGGTCAAGGTACACAGCGGGT 6397  
6307 CAGCCCTGTCAGAGCCCGCTGCTCAGACCGGCCCAAGGGGTCAAGGTACACAGCGGGT 6366  
6398 GTTCACTTGGCGCCAGCAATCAGTGAAGTATCAACAGACTTACACCGCGCAACCCA 6457  
6367 GTTCACTTGGCGCCAGCAATCAGTGAAGTATCAACAGACTTACACCGCGCAACCCA 6426  
6458 CAGCAGCTCAGCGACCCCTGCGCGCGCGCTTACTCTTCTTGGGAGCAGTCCCG 6517  
6427 CAGCAGCTCAGCGACCCCTGCGCGCGCGCTTACTCTTCTTGGGAGCAGTCCCG 6486  
6518 GTCCTGGAACCTCGCGCGCCACCACTGACTTACTCTCGCGCGCGCGCACTGCTGCC 6577  
6487 GTCCTGGAACCTCGCGCGCCACCACTGACTTACTCTCGCGCGCGCGCACTGCTGCC 6546  
6578 CCGGCGCGGTGCTCCCGCCACAGCGAAGGGGCGCAAGGTTCTCAGAGCCAAACAGAG 6637  
6547 CCGGCGCGGTGCTCCCGCCACAGCGAAGGGGCGCAAGGTTCTCAGAGCCAAACAGAG 6606  
6638 TCGGCTTGGGTGGTGGTGAAGACGGTATTGAACTGTGTCCCGACCGGAGGGCATGAG 6697  
6607 TCGGCTTGGGTGGTGGTGAAGACGGTATTGAACTGTGTCCCGACCGGAGGGCATGAG 6666  
6698 GAGCCAGGCGACTCCCGAGGTGCTGTATCCGCTGCTGACCGGATGGGAAACAGAG 6757  
6667 GAGCCAGGCGACTCCCGAGGTGCTGTATCCGCTGCTGACCGGATGGGAAACAGAG 6726  
6758 GAGCCAGGCGATGGGCTTCAAGTCTCAAGGCAACAGCAGCGCGCTTCTTTC 6817

Db 6727 GAGCCAGGAGATGGGCTTCAAGTCTCAAGGCAACAGCCAGCGCGCGCTTCTTC 6786  
6818 AGCAAGCTGACCGAGAGCACTCCGCTATGTCAAGTCCAAAGAGCAAGATCAACAG 6877  
6787 AGCAAGCTGACCGAGAGCACTCCGCTATGTCAAGTCCAAAGAGCAAGATCAACAG 6846  
6878 AAGCTGAAACCCCAACACCGGAAATGAGCTGAATACATATACGACCGCTGGGACGAG 6937  
6847 AAGCTGAAACCCCAACACCGGAAATGAGCTGAATACATATACGACCGCTGGGACGAG 6906  
6938 ATCTTCAATATGCGCGCTATCACCGGAAACAGGCTTATGACCTATAGAGCGCGGT 6997  
6907 ATCTTCAATATGCGCGCTATCACCGGAAACAGGCTTATGACCTATAGAGCGCGGT 6966  
6998 CAGGAACATGCGCAGCAACATGAGGGCTGAGGCGCATATTAAGAAAGCATCATGGT 7057  
6967 CAGGAACATGCGCAGCAACATGAGGGCTGAGGCGCATATTAAGAAAGCATCATGGT 7026  
7058 AATATGACCAAGTGGAGAGTCCCGCGCTCAGCGCAATGCTTTTAACTTGAAT 7117  
7027 AATATGACCAAGTGGAGAGTCCCGCGCTCAGCGCAATGCTTTTAACTTGAAT 7086  
7118 GCCAGTGCAGCGCTCCCGCTGATGCTATGCTATGCTGCTGAGCGAGAGTACAC 7177  
7087 GCCAGTGCAGCGCTCCCGCTGATGCTATGCTATGCTGCTGAGCGAGAGTACAC 7146  
7178 ACACCTACCTCGCAGGTGGCGCGGGAAGGCAAGGCTCTGAGGAGACCGCAGCCGA 7237  
7147 ACACCTACCTCGCAGGTGGCGCGGGAAGGCAAGGCTCTGAGGAGACCGCAGCCGA 7206  
7238 AAGCCAGATCCCGCGCGCGCGCTGAGCTGAGGAGACCGGCAACCTTGTCTCTCA 7297  
7207 AAGCCAGATCCCGCGCGCGCGCTGAGCTGAGGAGACCGGCAACCTTGTCTCTCA 7266  
7298 GTGCACTGGAAGAGACTGCAACCGCGGAGCGCGCTCAACCAACCGCTGTGGAGAGC 7357  
7267 GTGCACTGGAAGAGACTGCAACCGCGGAGCGCGCTCAACCAACCGCTGTGGAGAGC 7326  
7358 AAGCCCTGTCGCGAGGTTCAGAGCCATCCCTTACACCCCTGATCAAGCGGCTGAG 7417  
7327 AAGCCCTGTCGCGAGGTTCAGAGCCATCCCTTACACCCCTGATCAAGCGGCTGAG 7386  
7418 GCGGTGTATGAGCTTCCCGACCGCGCGCTTCCCGCGGAGCAGCGGCGCTTGT 7477  
7387 GCGGTGTATGAGCTTCCCGACCGCGCGCTTCCCGCGGAGCAGCGGCGCTTGT 7446  
7478 GCGCCCGACACGCTTGGAGCAGAGGACCAAGCACTGCTGTGCTGCAATACAGACA 7537  
7447 GCGCCCGACACGCTTGGAGCAGAGGACCAAGCACTGCTGTGCTGCAATACAGACA 7506  
Qy 7538 CTCTCCGACAGCGAG 7552  
Db 7507 CTCTCCGACAGCGAG 7521

RESULT 10  
US-09-819-104A-4  
; Sequence 4, Application US/09819104A  
; Publication No. US20030027137A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, J. Don  
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COMPRESSION MOLECULES  
; TITLE OF INVENTION: AND USES THEREFOR  
; FILE REFERENCE: UMG-030  
; CURRENT APPLICATION NUMBER: US/09/819,104A  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: 60/193,138  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 8544  
; TYPE: DNA

ASN

ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (160)..(7545)  
US-09-819-104A-4

Query Match 58.8%; Score 5037.2; DB 3; Length 8544;  
Best Local Similarity 78.3%; Pred. No. 0;  
Matches 6770; Conservative 0; Mismatches 1493; Indels 387; Gaps 45;

1 CATGTGGGCTCCACAGAGCTTGTGGCAAGACGTGGAGGGGCATGTGAGCCCGCTACCC 60  
159 CATGTCAAGATCAACAGAGCTGTGGCAAGACGTGGAGGGGCATGTGAGCCCGCTACCC 218  
61 GCCCACAAGCCCTTCTCAATCCAGTCAATCCCGGACGCAACAGAGCTGGGCTCT 120  
219 ACCCATGGCAATCTCTACCCCGGTGCAAGTACCCGGTCCCAAGAGAGTGGGCTGT 278  
121 GAGATCAAGACCACTCCCGGACTATGCTCCCACTGTGCGCGGCTCATATCCA 180  
279 TGAGTACCAACACACCCCGGTACTACACTCACACTGTACCCGGTTCATATCCA 338  
181 GCCCAGGCGGAGAGCCCTCCCTGTGTGTGAGTTCCAGCCCGGGAATGAACGGTCCA 240  
339 GCCACAGAGAGGCGGCGCTCACTGTGTGAGTTCCAGCTGGAGAGTGAACGGTCTCA 398  
241 GAGCTCCACCTGCGGACAGAGTCCCACTCATCTGCGCGAGCTGGGGAATCAGAGT 300  
399 GAGGCTTCAACTGCGCCCTGTGAGTCCCGACGTTCTGTCTGAGCTGGGCAAGCCGACAT 458  
301 GAGATTCAATTGAAGCAAGCGCCCTGCGCTAGAGCTGTGCTGTGACCCCTGTGCGACC 360  
459 AGAATTCAACGAGACAAAGCGCCCGCTGAGGCTAACCCGATACCTGTGCGGCC 518  
361 GTACCCCTGTGAGCAAGGCGCAAGCTGCGGGAATCTGAAGACTCAACCAAGAACCGTAG 420  
519 ATCAACCTGTGTGCTCACTGGGACCGCAAGTGGGTGAAGACTTACCAAGAACCGTAG 578  
421 CTTGACGGGCAAGCTGGAACCGGTGTCTCCCGCAGCCCGGCAACTGAACCTGAGCT 480  
579 CTTGGCAGGCAAGCTGTGAGCTGTGTCTCACTCCAGTCCCGGCAAGCTGAGCTGAGCT 638  
481 GAGCTGTGCTCCGCAAGCTGTCTCAAGAGAGAGCTGATCCAGAAC--TGAACCGCT 537  
639 AGAGCTGCGCGCATCTCACTGTCCAAAGAGAGCTGATCCAGAACAGATTGAGACGCGCT 698  
538 GGAACGGAAGATCAACATGTGTAGACAGAGATCTCTAAGCTGAAGAGAACACACACA 597  
699 GAACTGTAGATCACTATGTAGAGCAGAGATCTTCAAGCTGAAGAGAACACACACA 758  
598 GCTGAGAGAGAGAGCTGCAAGCGCCGAGGCTGTGAAGACCGGTGTCAACGCGCCAT 657  
759 GTTGAAGAGAGAGGCGCCGAGGCGCCGAGACCGAGAGCTGTGTGCGCACACCAT 818  
658 CGAGTGAAGCAGCGCAGCTGTGTGAGATCTATCAAGCAGAGAACCGAGAGAGGCTGA 717  
819 AGAATCAAGCAGCGAGAGCTGTGTCAAGATCATCTACGATGAGAACCGAGAGAGCGGA 878  
718 AGCTGCAATCGGATTTGTGAAGGCTGTGGGCGCCAGAGTGTGAGCTGTGCTGTACACA 777  
879 AGCCGCAACCGGATCTTAGAAGGCTGTGGGCGCCAGAGTGTGAGCTGTGTGTACACA 938  
778 GCCCTCCGACACCGGAGATCATGAGAACATCAAAATTAACCGAGCGATCCGAGAGA 837  
939 GCCGCTGACACACGCGAGTACATGAATAATCAAAATTAACCGAGCGATCCGAGAGA 998  
838 GCTAATCTTTGTACTTCAAGAGAGAGATCAAGCTGTGAGAACATGAGCAGAGATTCTG 897  
999 GCTAATCTTTGTACTTAAAGCGAGAGACACGCGGCAAGAGTGTGAGAACAGCGCTTCTG 1058  
898 CCAAGGCTATGACAGAGCTCTAGAGGCTTGGAAAAAAGGTGAGCGCATGAAAAACA 957  
1059 CCAAGGCTATGACAGCTCTAGAGGCGTGGAGAGAGAGTGAAGCGCATGAGAACAA 1118

958 CCGCGCCGCGCGGCAAGAGAGCAAGGTGCGGAGTACTACGAAGAAAGATTCCCTGA 1017  
1119 TCCGGAAGAGAGGCGCAAGAGAGCAAGTGTAGGGAGTACTACGAAGAAAGATTCCCGGA 1178  
1018 GATCCGCAAGAGCGCGAGCTGTGAGAGCGCATGTGAGAGCAGGGTGTGGCCAGCGGGCGAG 1077  
1179 GATCCGCAAGAGCGCGGAGCTGTGAGAGCGCATGTGAGAGCAGGGTGTGGCCAGCGGGCAG 1238  
1078 TGGGCTGTCCATGTGTGCGCGCGCGGAGAGACAGAGGTTCAAGATCATGTAGGCTT 1137  
1239 TGGGCTGTCCATGTGTGCGCGCGCGGAGAGTGAAGGTTTCTGAGATCATGTAGGCTT 1298  
1138 CTCAAGAGAGAGAACTTGAAGAACAGATGCGCCAGCTGTGCTGTGATCCCGCCCATCT 1197  
1299 GTCTGAGAGAGAACTTGAAGAACAGATGCGCCAGCTGTGCTGTGATCCCGCCCATCT 1356  
1198 GTACGAGCTGTACAGAGCGCATTAAGTTCATCAACATGAACCGGACTTATGTGCGGACCC 1257  
1357 GTACGAGCGC-GAGCAGAGAGATTCAGTTTCATCAACATGAATGACATCATGTAGTACCC 1415  
1258 CATGAAGGTGTACAAAGACCGCAGGTATGAACATGTGAGTGAAGAGAGAGAGAC 1317  
1416 CATGAAGGTGTACAAAGACCGCAGGTATGAACATGTGAGTGAAGAGAGAGAGAGAC 1475  
1318 CTTCCGGAAGAGTTTCATGACAGCATCCCAAGAACTTTGGCTGTATCGCATATTCCTGGA 1377  
1476 CTTCCGGAAGAGTTTCATGACAGCATCCCAAGAACTTTGGCTGTATCGCATATTCCTGGA 1535  
1378 GAGGAAGCAGTGTGTGAGTGTCTCTATTTACTACTGACTGAAGAAATGAGAACTTA 1437  
1536 GAGAAAGCAGTGTGTGAGTGTCTCTATTTACTACTGACTGAAGAAATGAGAAATTA 1595  
1438 TAAAGCCTGTGTGAGAGAGAGCTATCGGCGCGCGGCAAGAGCCAGCAACACACA 1497  
1596 CAAGAGCTTGTGTGAGAGAGAGCTATCGGCGCGCGGCAAGAGCCAGCAACACACA 1643  
1498 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1557  
1644 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1703  
1558 GAAAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1617  
1704 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1763  
1618 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1677  
1764 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1823  
1678 CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1737  
1824 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1883  
1738 AGGCGCATCAACCGCTTAATGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1797  
1884 AGGCGCATCAACCGCTTAATGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1943  
1798 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1857  
1944 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2003  
1858 AATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1917  
2004 GATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2063  
1918 GATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1977  
2064 CATGTTGGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2123  
1978 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2037  
2124 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2183







Db 4189 CCACTCCCCCAACAACCTCGGGAAGCTGACTGAGACCTTACAAGCCCCGACCTTGAC 4248  
Qy 4205 ---CTGGGCCCCCTGAAAGCTGAAGCGGAGCCATGAAGGCTGTGTGGCAACGGTGAAGAG 4261  
Db 4249 CCTTGGGTCCTCTGAAAGCTGAAGCGGACTCACAGAGGATGTGTAGCAACTGTGAAGAG 4308  
Qy 4262 GCGGAGCCGCTTCATCATGAGATCCGCGAGAGAGCTGCGGCAACGCCCCAGCTGCCC 4321  
Db 4309 GCGGAGCCGCTTCATCATGAGATCCGAGAGAGAGCTGCGCGCAACTGTAGCTACCC 4368  
Qy 4322 CTGGCCCCGCGGCGCTCAAGAGGCTTCATCAACGAGGCAACCCCGCTCAATGACAC 4381  
Db 4369 CTGGACCAACGAGCTCTGAAGAGGATTCATCACCAAGGCAACCCCACTCAAGTACGAC 4428  
Qy 4382 ACCGCGAGCTCCACCACTGGCTCAAAAAGCAAGATAGCTGCTCCCTCATCGGACGCCC 4441  
Db 4429 TCTGGGAGCACTTCATCTGGCAACAGAAACAGAGTGGCTTCATCATCGGACGCCC 4488  
Qy 4442 GCGCGGAGCTTCCACCCGATGCAACCGCTGATGTGTAGCGAGCGCCCGGCACTGAA 4501  
Db 4489 GCGCGGAGCTTTCCTGGCTCTGACCGGCTGACATTAATGGCTGAGCGCCGGGCACTGGAG 4548  
Qy 4502 CGTGGCTCTTACAGAGAGGCTTGAAGAGCGGCGCAAGGAGACCGCCAGCACTCGGAGGAC 4561  
Db 4549 CGTGGCTCTTATGAAGAGAGTGTGAAGAGCGGCTCAGGAGCAGAGATGTGTCAAGGAGC 4608  
Qy 4562 TCCATTGCGCGGCGGCGCCCGGCTATGTGAGCTGAGCTGGATTAAGCGCGGAGAGCCCC 4621  
Db 4609 TCCATTCACAGTGGGAGCTCAAGTCTGTGCTTGAATGAGGCAAGGCAACGAAAGCCCA 4668  
Qy 4622 CTGACTATGAGAGCAACGAGGCAACCTTTGCGGCGCACTCCCAAGAGTTGCGCCGTG 4681  
Db 4669 CTGACTTACGAAGACCAAGGAGCACTTTCACAGTCACTGCGCAAGTGGCTCCCTGTG 4728  
Qy 4682 ACCATGCGGAGAGCCCAACCGCGGCTTGCAGAGAGGAGGCTTTGCTTCAGCAAGGATCC 4741  
Db 4729 ACCAAGAGGAGAGCCCAACCGCAAGGCTTTCAGAAAGGAGGCTCTTATCCAGCAAGGCTCC 4788  
Qy 4742 CAGGACCGAAGCTGACGTGCAAGCTGCTGTGAGATGAGCAAGTCCCGGCAAGCAACCGTG 4801  
Db 4789 CAGGACCGGAGAGCTGACATCTTACCCCGGAGATGAGCAAGTCCCGCAACAGCACTGTG 4848  
Qy 4802 CCGGAGCAACCAACCAACCCCATCTGCGCTTATGAGCACTGCTTGGGAGCTGAGTGGC 4861  
Db 4849 CCGGAGCAACCACTCAACCCCATCTGCTTATGAGCACTTGTGCGGGGCGTGAAGTGTG 4908  
Qy 4862 GTGAGCTGTATGCGAGCCCAATCCCGCTGAGCTTGCATCCCACTTCATCCCGCGGC 4921  
Db 4909 GTGAGCTGTATCGGTGTCAATCCCATTTGGCTTTGACCCCACTTCCATGCCGAGG 4968  
Qy 4922 ATCCCTGTGGA---GCGAGCGGCTGCTACTACTGCGCCCGCAACCTTGGCCCCCAACCCC 4978  
Db 4969 ATCCCTGTGGAAGAGAGAGCCGAGCTTACTTACTGCGCCCGGCACTTGGCCCCAGCCCC 5028  
Qy 4979 ACTTACCGGCACTGTATCCCACTCATCTGAGCTGCGGCTAACCCGCAACGAGCGGCTG 5038  
Db 5029 ACTTACCCACACTGTATCCCACTTATCTCATCCGAGGCTTACCCGAGCAACGAGGCGCTG 5088  
Qy 5039 GAGAACCGGAGACATCATCATATGACTTCACTTCCGACAGATGACCAACAGCT 5098  
Db 5089 GAGAACCGGAGACATCATCATATGACTTCACTTCCGACAGATGACCAACAGCT 5148  
Qy 5099 GCGACCGGAGAGCGGAGAGCTGATGTGAGGAGGCGCTGCGGCGGAGAGTCTG 5158  
Db 5149 GCTTCGCGCATGCGGAGGCTGTGACTGTAGGAGGCTGTGACCGGAGAGTCTG 5208  
Qy 5159 CTGGCACTCAACTGAGCTGCGGCTCCCGAGGACATCATGAGCTGTCCCAATGTGCCACAC 5218  
Db 5209 CTGGGCTCTAATTAATTCGCGTGGCCCAAGAGCATTAATGACTGTGCCAAGTGCACAC 5268  
Qy 5219 CTGGCTGTGTGTGCGCCCGCAACAGGAGCAACCCGAGCAACGAGCAACGAGCGCTTGTG 5278  
Db 5269 CTGGCGGTGTGTGCAACAGCGAGGCAACCCCTGCAACGAGCATGACGCGCTTGTG 5328

Qy 5279 TACCTCCCAACCGCGCCCGAGCCCTTTCAGAGCGGCAACAGAGCTCCCACTTCCCA 5338  
Db 5329 TACCTCCCAACCGCGCCCGAGCCCTTTCAGAGCGGCAACAGAGCTCCCACTTCCCA 5388  
Qy 5339 GAGAGTCCACACTTTCAGAAACCAACCAACAGCTCTGTCCGAGCGGAGAGAGAC 5398  
Db 5389 GAGAGCGGCACTTCACTTGAACCAACCTGCAATCTTTCAGAGCGGAGAGCG--- 5444  
Qy 5399 GCGGATTCAGAGCGGAGACCGGAGATCGGAGGAGGAGAAAGTCCATCTCAAGTCCACAG 5458  
Db 5445 -----GGAAGTATGCGGAGAGAGCAAGTCCATCTCAAGTCCACT 5490  
Qy 5459 ACAGTGAAGCAACGCACTTGAAGCTGTGAGAGCTGTGACAGAGAGAGCGGAGAGCGG 5518  
Db 5491 ACAGTGAAGCAACGCACTTGAAGAGCTGTGAGAGCTGTGAGAGAGAGAGCGGAGCGG 5550  
Qy 5519 AGCAGCGCGGAGGAGTGGGAGAGAGAGAGCGGCGCCCGCTCCCACTCCATGCGCACAG 5578  
Db 5551 AGCAGCGG-----CGCGCTCCCAACCAACAG 5580  
Qy 5579 CACTGCGCATCTCCCTCGGAGCCGAGATGCGCTCAAGAGAGACCAAGTGTGCTTAC 5638  
Db 5581 CACTGCGCATCTCCCTCGGAGCCGAGAGAGCTGTGAGAGAGAGAGAGAGTGTGCTGAC 5640  
Qy 5639 AACACAGGCAATGAAGGATATCATCACCGCTGTGAGAGCCAGCAACCCAGCTCTGAGG 5698  
Db 5641 AACACAGGATGAAGGAGTGTATCATCTCGTGAACCGGAGCAACCGGCTCTGAGG 5700  
Qy 5699 TCCACTTCACCTCTCTACACCGGTTGGCCAGCTGCAATTTCCACCTGCAACCACTGCG 5758  
Db 5701 TCCACTTCACCTCTCTGCTGTGCGCCAGCTGCAATTTCCACCTGCAACCACTGCG 5760  
Qy 5759 CCAGTGGGAGGAGCCCTGATGAGGAGTACCTTACCTCATGAGAGCGGCTTGTGCGCC 5818  
Db 5761 CCAGTGGGAGGAGCCCTTGAAGGAGTCTACCTTACCTCATGAGAGCGGCTTGTGAGCC 5820  
Qy 5819 AAGAGGCGCCCGGAGTGCAGCGGCGAGAGCGGCGGAGAGAGAGAGAGAGAGAGAGCTTC 5878  
Db 5821 AAGAGAGCTCTCGGAGTGCAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGCTTC 5880  
Qy 5879 CTGCGCAAGCCCGGAGCGGCTCGGAGCTGAGAGCGGCTCTCTCCCGAGCAAGGCTGCG 5938  
Db 5881 CTGCGCAAGCCCGGAGCGG-----GAGAGCGGCTCTCTCAACCGAGCAAGAGCTTC 5931  
Qy 5939 GAGCGCGGCGGCTTGTGCTGTCTGTCTGTGCGGCAAGGCAAGTGGCGGAGCGGCTGCG 5998  
Db 5932 GAGCGCGGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5991  
Qy 5999 AAGAGCTTGCAGCTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6058  
Db 5992 AAGAGAGCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6048  
Qy 6059 GAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6118  
Db 6049 GATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6108  
Qy 6119 TCTTGGGTATACCA---GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6175  
Db 6109 TCTTGGGTATACCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6168  
Qy 6176 GTGAGCTCAACGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6235  
Db 6169 GTGAGCTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6228  
Qy 6236 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6295  
Db 6229 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6288  
Qy 6236 GAGGCGGCTCACTTCCCACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6355  
Db 6289 GAGGCGGCTCACTTCCCACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6348

6356 CTGCTCAGAACCGCCCGAGGGGTCAAAAGTCAACAGCGGTGTGACCTTGCCCAAGAC 6415  
6349 CTCTTCAAGATGCCCCAGGCATCAAAAGTCAACAGAGGGTGTGACCTTGCCCAAGAC 6408  
6416 ATCAGTGAAGGTATCAACAGAGTACACCGGCAACCCCAAGAGCTCAGCGCACCC 6475  
6409 ATCAGAGAGGTCAATTACGAGAGCTACACCGGCAACCCCAAGAGCTCAGTGCCCC 6468  
6476 CTGCGCGCGCCCTCTACCTGCTTCCCTGGGGGAGAGCTGCGGCTGCTGAGCTCCGCGC 6535  
6469 CTTCGCGCCCTCTCTACCTCTTCCCGAGCCAGCTGCGCTGTGTGATCTTCGCGC 6528  
6536 CCAACCAAGTACCTCTACCTCCGCGCCGAGCAATGTGCCCCGCGCGCTGCTCCCC 6595  
6529 CCAACCAAGTACCTCTACCTCCGCGCCGAGCAATGTGCCCCGCGCGCTGCTCCCC 6588  
6596 CACAGCGAAAGGGGGCAAGAGTCTCAAGACCAACAGAGCTGTGGGTCTTGGGTGTGT 6655  
6589 CACAGTGAAGGGGGCAAAAGGTCCCAAGAACCAAGCAAAACATCGGTCTGGGCAAGAGT 6648  
6656 GAGGAGGTATTGAACCTGTGTCGCCACCGAGGGGATGACGAGCCAGGCACTCCCG 6715  
6649 GAGGATGCAATTGAGCTGTGTCCCAACAGAGGGATGACTGAGCCAGGCAATGCTCG 6708  
6716 AGTCTGTGTACCGGCTGTGTACCGGATGGGAAACAGACGAGCCCAAGATGGGC 6775  
6709 AGGCTGTGTACCACTGTGTATCGAGACGGGAAACAGGCGAGCC--CAGGATGGGC 6765  
6776 TCCAACTCTCAGGCAACACAGCCAGCGGCAAGCTTCTTCAGCAAGCTGACGAGC 6835  
6766 TCTAACTCTCAGGCAACACAGCCAGCGGCAAGCTTCTTCAGTAACTGATGAGAGC 6825  
6836 AACTCCGCCATGGTCAAGTCAAGAGCAAGATATCAAGAAAGCTGAAACACCCCAAC 6895  
6826 AACTCCGCCATGGTGAAGTGAAGAGCAGAGATCAACAAAGACTCAACACCCCAAC 6885  
6896 CGGAATGAGCCTGTAAATATATACAGCAGCTGGGACGGAGTCTTCAATATAGCCGCG 6955  
6886 CGGAAGAGCCAGAAATACAAATATTTGGCAGCTGGGACGGAAATCTTCAACATGCGCGC 6945  
6956 ATCAACCGGAACAGGCTTATGACTTATAGAGCCAGGCGGTGAGGAACATGCGACAC 7015  
6946 ATCACTGAGAGAGGCTTATGACTGTAGAGCCAGGCGGTGAGGAACATGCGACAC 7005  
7016 AACATGGGGCTGAGAGCCATATTTAGAAAGGCACTATGGGTAAATATGACATGTGGAA 7075  
7006 AACATGGGGCTAGAGGCAATTTATTAAGGCACTATGGGTAAATATGATCATGTGGAA 7065  
7076 GAGTCCCGCGGCTCAGCGCAATGCTTTAACCTCTGAATGCAAGTGCAGCTGCGC 7135  
7066 GAGCCCCCGGCTCGGCGGCAATGCTTTTAACTCTGAATGCAAGCGCAAGTGTGCC 7125  
7136 ---GCTGTATGCGCATTAACCGCTGTGACGAGCGAGTGAACACATCACTGCGCA 7192  
7126 GCTGTGTATGCGCATTAACATGCTGACGAGCGAGTGAACACGCACTCACTGCGCA 7185  
7193 GGTGTGCGGGGAAGGCAAGGTCTCTGGAGACCCAGACCGGAAAGCAAGTCCCG 7252  
7186 GGTGTGAGTGGGAAACCAAGGTCTCTGGAGACCTTAGCAGCGGAAAGCAAGTGGCA 7245  
7253 GCGCGGGGCTGGCATCTGGGAGACGGGCAACCTGTGTCTCAAGTGACTGAGGGAG 7312  
7246 GACACAGGCTTACGTTCGAGAGCCAGCCCTTCTGTCTCTCAAGTACATCAAGAGGG 7305  
7313 GACTGCAACCGCGGACGCGGCTCAACCAACGCGTGTGGAGGACAGGCGCTGTGCGCA 7372  
7306 GACTGCAATCGCGGAACACCACTCAACCAACGCTGTGGAGGACCGGCGCTCATGCGCA 7365  
7373 GGTTCAGAGCATTTCCCTTCAACACCCCTGTATATGCGGTGAGAGCGGTGTCAAGGT 7432  
7366 GGTGTACAGCATTTCCCTTCAACACCCCTTGTATATGAGGCTTAGAGCAAGGTGTCAAGGT 7425  
7433 TCCCAACCCCAACGGGCGCTCCCGGGGAGAGCGGCGCTCGCTGAGCCCAACGCGC 7492

7426 TCCCGCGCCCACTGAGCTTGGCGGAGGAGCGGCGCCCTTACCTGATCCCAACGCGC 7485  
7493 TGGAGAGAGAGCCCAAGCACTGCTGTGCAATGAGAGCACTGTCCGACAGCGAG 7552  
7486 TGGAGATGAGAGAGCCCAAGCACTGCTGTGTTACAGATGAGAGCACTCTGAGAGGAG 7545  
7553 TGACTCAGAAAGAGCGGGGGGGGGCGGCGGTGEBURPERLUTGAGGTCCAGCGAGC 7612  
7546 TGAACAGGATTTGGGGGGAGCGG-----TGCAAGGTCCCGCAAA 7586  
7613 CACAGGAACGGCTCTGACAGAGCGGGGCGGCTGCCAATCCCAACAAAGAGAGGCC 7672  
7587 GCGAAGAGAGCCCAAGCACTGAGCAAGACTGTCACTCGGAGCTGAGGAAGAGGCC 7646  
7673 CCTGAATCCGCGTC-----GCTTCATCATCTGTCCGTCCAGAGCCGG 7717  
7647 CTGAAGTGTGCTGTGCTGTCATTCGTCCTGCTCACTATCTGTCTCAATCCAGAGCTGG 7706  
7718 CATCTTGTCTGTAAAGCTTAACTAAGACTCCGCGCGGCGCTGCGCTGTGAG-- 7775  
7707 CATCC--TGCTGTCTTAAGCTTAACTAAGACTCCACCCCGGCTGCGCTGTGAGT 7765  
7776 --ACTTACTCAGGGAGTGTTACTGTGTCTGAGAGGGAGAGAA----- 7821  
7766 ACTTACACTCAGGGAGTGTTACTGTGTCTGAGAGGGAGTGAACAGAGGGAGAGG 7825  
7822 ----GGGCGCGGAGGGGGGACCGGAGCGGTGTGAGCCACACACAGCGGCGAGG 7876  
7826 ACAGCGGGGCAAGAGGGGGGGGACAGCAATCGTGTCAAGTGCACATCGTGTCTGGGG 7885  
7877 CGGCGAGGAGCCCAAGAGAGTATGACAGCACTTCAGCGCACTGCT--CCCGCAAT 7934  
7886 TCAAGGGAGACCAACAGAGCTGACAGGCACTTCAAGCAAGCTGTGCGCTTACC 7945  
7935 GCATTGGAACCAAGTCTTAACTGAGCTGACAGCCCGCGGCTCTCGCTTCCA 7994  
7946 GATTTGGAACCAAGTCTTAACTGAGACTCTGCGGTGATCTGCGCTCTGTGCGCCCA 8005  
7995 TCCCGCTTACGCTCTGAGCAATGAGCCGAGGCTGTTCAGAGCCCGAGTGGCTGTT 8054  
8006 GCGCGCTTCTCTGTGAGCAAGAGCTTCCAGCTTATCTGCTTAACTGTGATCA 8065  
8055 CGGCTCCCAAGACTGCGCCAGCGCAAGAGATGCTGGAACCAAGTCAAGCGAGGTG 8114  
8066 TCGAGTCTTCAAA--GGCAGCCAGCCCAAGAGCTGGAGCCCATCAAGACAGGTG 8122  
8115 GCGGACAAAGGGCCAGGTGCGGCTGAGGGGAAACGAGATGCTTCCAGAGACTGACTGTT 8174  
8123 GTGACAAAGGGGCGTGG-----CTGGGGCAAGGATGCTTGAAGAACTGACCGTT 8174  
8175 TTTTACACATGTTGCGCGAGCGGTGGAGAAAGGCAAGTATTAATGATGTGTTGT 8234  
8175 TCCCGGCTGTGTGTGTGCAACGCG---GAGGGAAGGACGTTGTAATG--GTGTTGGC 8228  
8235 TTAAGGGTAAATTTTGAATCTTCAATGAATTAATTCAGATGTTTTCGCAAGGAGG 8294  
8229 TTAAGGGTAAATTTTGAATCTTCAATGAATTAATTCAGAGTCTCAACAAAGAGG 8288  
8295 ACTTACCAAGTATTAAGTGTGTGCTTGTGATCTCTGCTTAACTGTTCAAGAGGCTGT 8354  
8289 ACTGCGCC--GTGTTCTCCGCTGTGC--TTTGTGTCTAAGCTTATGTTCAAGAGCGA 8346  
8355 GCAAGCCGACAGTGGTGAACCCCATATCTGCGAGACCAAGGGGGGGGAGCTGTCTGC 8414  
8347 GC---CAGCCAAAGGCGGTGGCCCATACGAGAGCTTGGGGGTGAGGGGCTCCGAGAC 8403  
8415 AGCGCCCGGTGTGCTTCCCTCCCTCCCTTCTTGGGAGAAATGAATGAATGATGAGTATTC 8474  
8404 AGGCACTGTG--CCCTTCCCAACCTTACTTCAAGCAAAATGAGCTGATGTGATTC 8460  
8475 TGTGCGGCAATTTGGAGAGGGTGTGATTTCTGTCTATTTACACAGTGTCTTAATTA 8534

```

Db      8461 TGTGGCGCCTCTCTGTGCA---CGGCGGCATTCTGTATTACACATGTTGTTCCAAATTA 8517
Oy      8535 AAAAGCGCAT 8544
        |||||
        |||||
Db      8518 AAAAGCAAT 8527

RESULT 11
US-10-764-420-1526
; Sequence 1526, Application US/10764420
; Publication No. US20050084872A1
; GENERAL INFORMATION:
; APPLICANT: Lum, Pek Yee
; APPLICANT: Tan, Yejun
APPLICANT: Dai, Hongyue
TITLE OF INVENTION: Methods For Determining Whether An Agent
TITLE OF INVENTION: Possesses A Defined Biological Activity
FILE REFERENCE: ROSA122057
CURRENT APPLICATION NUMBER: US/10/764,420
CURRENT FILING DATE: 2004-01-23
PRIOR APPLICATION NUMBER: US 60/442,797
PRIOR FILING DATE: 2003-01-24
PRIOR APPLICATION NUMBER: US 60/474,413
PRIOR FILING DATE: 2003-05-30
NUMBER OF SEQ ID NOS: 3683
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1526
LENGTH: 8544
TYPE: DNA
; ORGANISM: Mus musculus
US-10-764-420-1526

```

Query Match	58.8%	Score 5037.2	DB 9	Length 8544
Best Local Similarity	78.3%	Pred. No. 0	Mismatches 1493	Indels 387
Matches 6770	Conservative	0	Gaps 45	
QY 1	CATGTCGGGGCTCCACACAGACTTGTGTGGCACAGACGTCGTGAAGGGCCACTGAGCCCGCTACCC	60		
Db 159	CATGTCAGGATTCACACAGCCTGTGTGGCACAGACATGAGGGGGCTGTGAGCCCGCTACCC	218		
QY 61	GCCCCACAGCCTTTCTTACCAAGTCAGATGCGCCGACGACACGACGTCGGGCTCTT	120		
Db 219	ACCCCATGGACTTCTCTACCCGGGTGCAATAGCCGGTCCACACGAGCGTGGGGCTGCT	278		
QY 121	GGAGTACAGAGACCACTCCCGCGACTAAGCTCCCACTGTTCGCGGGCTCATCATCA	180		
Db 279	TGAGTACCAACCAACCCCGGATCACTCACTCACTGTCAACCTGTCAACCTGGTCCATATCA	338		
QY 181	GCCCCAGCGGCGAGAGCCCTCCCTGCTCTTGAATTCAGCCCGGGAATGAACGGTCCA	240		
Db 339	GCCACAGAGAGGCGGGCCCTCACTGTCTTCAGAGTTCCAGCTTGGGAATGAACGGTCTCA	398		
QY 241	GGAGCTCCACTTGGGGCCAGAGTCCACTCATCTGCCCCGAGCTGGGAAATCAGAGT	300		
Db 399	GGACTCCACCTGGCGCCTGAGTCCCGACGTTCTGCTGAGCTGGGCAAGCCGACAT	458		
QY 301	GGAGTTCATTAAAGCAAGCGCCCTCGGCTAAGAGTGTGCTGACCCCTCTGGCAC	360		
Db 459	AGAAATTCACCGAGAGCAAGCGCCCCCGCTGAGTACTACCCGATCCCTGTGCGCCC	518		
QY 361	GTCAACCCCTGTGTGGCACAGGGCACGCTGTGGGATCTGAAGCTTCACCAAGAACCTTAG	420		
Db 519	ATCAACCCCTGTGTGGCACCTGGGCAAGCCAGTGGGTCTGAAGCTTCACCAAGAACCTTAG	578		
QY 421	CCTGACGGGCAAGCTGGAACCGGTGTCTCCCCAGCCCCCGGCACTGACCTTGAGCT	480		
Db 579	CCTGGCAAGCAAGCTGAGCCTGTGTACTCTCCAGTCCCGGACGCTGACCTTGAGCT	638		
QY 481	GGAGCTGTGTGCGCACAGGCTGTCCAAAGAGAGCTGATCCAGAAC---TGCACCGGCT	537		
Db 639	AGAGCTGGCGCATCTTCAGACTGTCCAAAGAGAGGTGATCCAGAACCAATTGGACCGGCT	698		
QY 538	GGACCGAGATATCACCATGTATGAGACAGACAGATCTCTAAGCTGAAGAGAAAGCAGCA	597		

[illegible]

Db 1764 GAAAGAAAGAAAGAACTCAGCAAGAGAAAGACAGACGACTTCTGGCGAGGACAAACGA 1823  
Qy 1678 CGAAGAGAGGCTGTGGCTCCAAAGCCGCAAACTGCGCAACAGCCAGGAGAGACGCA 1737  
Db 1824 TGAAGAAAGAGGCGGTGGCTTCCAAAGCCGCAAACTGCGCAACAGCCAGGAGAGACGCA 1883  
Qy 1738 AGGCGGCATCAACCGCTCAATGGCTAATGAGCCAAACAGGAGAGAGCCATCAACCCCA 1797  
Db 1884 AGGCGGTATCAAGCGCTCCATGGCCAAAGAGCCAAACATGAGAGAGACAGCCACCA 1943  
Qy 1798 GCAAGAGCGCGAGCTGGCTCCATGAGAGCTGAATGAGAGTTCCTGCTGGAAGAGAGAA 1857  
Db 1944 GCAAAATTGAGAGCTGGCTTCCATGAGAGATGAAAGAGAGTTCTGCTGGAAGAGAGAA 2003  
Qy 1858 AATGGAACAGCCAGAAAGAGTCTCCGGAACAGGCGCGCACTGGTCGGCCATCGCCG 1917  
Db 2004 GATTGAGACAGCAAGAAAGGCTCTTGGAACATGGAGAGAACTGGTCAGCCATTGGCCG 2063  
Qy 1918 GATGTGGGCTCCAGACTGTGTGCGAGTGTAGAACTTCTACTTCACTTCAAGAAAG 1977  
Db 2064 CATGGTGGCTCCAAAGCCGTGTCCAGTGTAGAACTTCTCACTTCAAGAAAGAG 2123  
Qy 1978 GCAAACTTCGATGAGATCTTGGACAGCAAGCTGGAAGATGAGAGAGAGAAACG 2037  
Db 2124 GCAGAACTTGAAGCAAAATCTTCAAGACAGCAAGCTTAAAGATGAGAGAGAGAAACG 2183  
Qy 2038 GGGGAAGAAAGAAAGAGCGCGCGCGCGCAGCGAGGAGGCTGCAATCCGCGCGT 2097  
Db 2184 TCGAGAGAAAGAAAGAAAGACCCAGCTGGCGAGCGAGAGAGAGACCTTCCACCTGC 2243  
Qy 2098 GGTGAGAGATGAGAGATGAGAGGCTGCGGCGTGAACGGAATGAGAGAGAGATGATGA 2157  
Db 2244 CGCTGAGAGAGAGAGATGAGAGATCAAGCGCAATGCGCAATGAGAGAGAGCTGGCGGA 2303  
Qy 2158 GGAAGCTGAAGCTTATCATGCTCTTGGGAATGAGGTGCCAGA--GGGAATGCAAGTGG 2214  
Db 2304 GGAAGGAGAAAGCTTCAAGGCTCTGGGAATGAGGTGCCAGAGTTGGGGAAGTGCAGTGG 2363  
Qy 2215 CCGAGGCACTGTCAAAAGAGCTCAAGACCCGAGACATCCCTCTCTCAACATGAGGC 2274  
Db 2364 CCGAGTGTGTCAAAAGAGCTGTGATCTGAGAGTGTCCCAATCCCGGTTCAAGAG 2423  
Qy 2275 CGCAGAGACAGAGGCGAGATGGGCCCAAGCCCCAGCCACCTGGGCGCGACAGGCG 2334  
Db 2424 CAGAGAGAGAC-----TGGGCTTAAACCCATCTGGGCACTGAAGCATTTGCCCGCTGC 2474  
Qy 2335 ACCCCAGGCGCCACCCACCAACAGAGAGCATCCGAGCCGCCATTGAGCCACCC 2394  
Db 2475 CACCCAGCCACTGTCTCTCTCAAGAAAGACCGGACAGCCCTGTGAGGCTCTCCG 2534  
Qy 2395 GGCCTTGAAGCCACCGGAGCCCTTACGCCCCCAAGACACCCCATTCGCCCTCTGACCC 2454  
Db 2535 AGTCCCTGATGCA---GTGGCCCAACATCCCAAGGCTTCCCACTCACTGCCACCC 2591  
Qy 2455 TCCTCTCTGTGTCCCAAGAGAGAGAGAGAGAGACCGACAGACGCGCCCACTGAG 2514  
Db 2592 CCGGCTTACTGTGAGAGAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2651  
Qy 2515 GGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2556  
Db 2652 AGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2711  
Qy 2557 AGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2616  
Db 2712 CAGAGGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2771  
Qy 2617 GGGGCGGCGC---CAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2673  
Db 2772 AGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2831  
Qy 2674 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2733  
Db 2832 TAAAG---GTGAGAGAGAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2888

Qy 2734 CCCCCAGAGACAGGACTTCAGATGCTTACCTGCAGTGCAGAGAGAGTGAAGCCGAGAG 2793  
Db 2889 CACCCAGAGACAGTGAATCTCAGTGCACCTGCAGTGCAGATGAGAGTGAAGAACCCGAG 2948  
Qy 2794 CCGGCAAGAAAGCCGGCTGTCTTCCCAAGGCTCCAGCTTCTCAACCCCACTGGCCACCC 2853  
Db 2949 AGGTGAAGAGGAGAGAGCTGTCTTCAAGAGGCTCCAGCTTCTCAACCCCACTGGAGATCC 3008  
Qy 2854 CCGGCGCAATGCTCAACCCAGAGGCACTGAGCTTGAAGAGAGTGAAGAGAGAGAGAGAG 2913  
Db 3009 CCGGCGCAATGCTTCCGCCAGAGAGGCTGTGAAGCTGAAGAGAGAGAGAGAGAGAGAG 3068  
Qy 2914 TGCATCCCCCATTCAGATCAAGAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2973  
Db 3069 CGGCAATCCCCCATTC---GTACAGAGAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAG 3125  
Qy 2974 CACCAAGCAGCTCCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3033  
Db 3126 CCAAGAGCAGATTCCTGTGCTTCAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3185  
Qy 3034 CCTCAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3093  
Db 3186 GTCTCAGAGTGGAGAGAGAGTCAAGTGCAGAGTCCGAGAGAGAGAGAGAGAGAGAG 3245  
Qy 3094 CAGAGAGC-----CTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3130  
Db 3246 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3305  
Qy 3131 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3190  
Db 3306 ACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
Qy 3191 AAGGCTCCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3250  
Db 3361 AAGACTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3414  
Qy 3251 CCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3310  
Db 3415 CTGCTTGGGCTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3471  
Qy 3311 AACCGGCTCCCTCATCTCTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3370  
Db 3472 AACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3531  
Qy 3371 GGCATTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3430  
Db 3532 GGCATTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3588  
Qy 3431 CCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3490  
Db 3589 CCGATGGGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3642  
Qy 3491 TTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3550  
Db 3643 ----- 3642  
Qy 3551 CTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3610  
Db 3643 -----GGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3662  
Qy 3611 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3670  
Db 3663 ACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3715  
Qy 3671 ACATACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3730  
Db 3716 AGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3774  
Qy 3731 ACAGAGATCATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3790  
Db 3775 AGCAGAGATGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3834

OY	3791	CCCAAGGCGCACGTATCTTAGAAGGACAAAGAGGCGCACGTCTTGTCTTATGAGGGTGGC	3853
Db	3835	CCCAAGGCGCACGTATCTTAGAAGGACAAAGAGGCGCACGTCTTATGAGGGTGGT	3894
OY	3851	ATGTCGTGACCCAGTGTCTCCAAAGGAGACGGCAGAACAGCTCAGACCCGCCATGAG	3910
Db	3895	ATGTCGTGTACAGTGTCTTAGAGAGATGGAAGAGACGTGGGGCCACCCCATAG	3955
OY	3911	ACGGCCGCCCCCAAGCGCACCTATGACATGATGAGGGGCGCGTGGGAGAGCCATCTCC	3970
Db	3955	ACTGCGCGCCCTTAACGCACCTATGACATGATGAGGGGCGCGTGTGAGGAGCTGTCAAC	4014
OY	3971	TCAGCGACATTCGAAGGTCTCATGAGGCCGTGTGCATCTCCGCGAGACCAACGCCCAAC	4030
Db	4015	TCAGCGACATTAAGAGGACTCATGAGGCGCGCCATCTCTGACAGC-----ACAGCCCC	4068
OY	4031	CACCTCAAAAGACAGACCAACATCCGCGGAGGATCATCAACAAGAGATCCCTCGGCTCAC	4099
Db	4069	CACCTCAAGAGACAGATCATCATCCGAGGCTTCATCAGCAAGAGCATCCGAGGTCTTAT	4128
OY	4091	GTCGAGGACAGAGAGACTACCTGCGTCGAGAGGCGCAAGCTCTTAAGCGGAGGGACG	4150
Db	4129	GTCGAGGCGCAGAGAGACTACTTACGCGGGAGGCGCAAGCTCTTGAACGAGAGGACA	4188
OY	4151	CCTCGCGCCCCACCGCCCTCACAGGAGCTTGACCGAGGCTTACAGACGACAGCC-----	4204
Db	4189	CCACTCTCCCCACCAACCATCTCGGAGCTGTGACTGAGCTTACAAACCCCGGCCCTGTGAC	4248
OY	4205	---CTGGGCGCCCTGAAGCTGAAGCGGCGCCCATGAAGGCGCTGTGTGGACAGGTAAGAG	4261
Db	4249	CCTCTGGGTCCTCTGAGCTTAGGCGCACTCACGAGGGTGTGTAGCACTGTGAAGAG	4308
OY	4262	GCGGGCGCTCCATCATGAGATCCCGCGCAGAGAGCTGCGGACACGCGCGAGCTGTCC	4321
Db	4309	GCGGGCGCTCTATCATGAGATCCCGAGAGAGAGCTGTGCGCGGACACCTGAGGTATCC	4368
OY	4322	CTGCGCCCGCGCGCGCTCAAGAGGGCTTCATCACGACAGGGCACCCGCTCAATGATGAC	4381
Db	4369	CTGGACACACGGCTCTTGAAGAGGGTTTCATCACCCAGGGCACCCCACTCAAGTACAC	4428
OY	4382	ACCGGCGGTCACACCATGGGCTCCAAAAGACAGAGTACGTCCTCATGGGACGGCC	4441
Db	4429	TCCTGGGGACCTCTCATCTGGACCAAGAAACAGACGTGGCTCATATGAGGACGGCC	4488
OY	4442	GGCGGGAAGTTCCACCCCGTGCACCCGCTGATGTGATGTGCGACGCGCCGGGCACTGGAA	4501
Db	4489	GGCGGCGCTTTCCTCGCCCTGACCCGCTGGACCAATATGGCTGACGCCCGGGCACTGGAG	4548
OY	4502	CGTGCCTGTCTACAGAGAGAGCTTGAAAGCGCGGCGCAGACCCGACGAGCTCGGGGGG	4561
Db	4549	CGTGCCTGTCTTGAAGAGAGTCTGAAAGCGCGGTCAAGGAAACAGAGTGTGACGGGGG	4608
OY	4562	TTCATTTGGCGCGCGCGCGGCTCATTTGTGCTGAGCTGGGTGAAGCGCGGCAAGAGCCCC	4621
Db	4609	TTCATTCACGTTGGGGCTTCAGTGTCTGTGCTTGACTGGGAAACCAACGCAAAAGCCCA	4668
OY	4622	CTGACTTATGAGACCAACGGGGCACCTTTTCCGGCGCACTTCCACGAGTGTCCGCGGTG	4681
Db	4669	CTGACTTACGAAGACCAACGGGGCACCTTTCACACAGTACCTGGCCAGTGGCTCCCTGTG	4728
OY	4682	ACCATGCGGAGAGCCACGCGCGCGCTGACAGAGGGGACGCTTTTGTCTCAGCAAGGACTCC	4741
Db	4729	ACCACGAAGGAGCCACGCGCACGCTTTCAGGAAGGACGCTTATTCAGCAAGGCGTCC	4788
OY	4742	CAGAGACGGAAGCTGACGTCGACGCGCTGTGAGATCGGCAAGTCCCGGCAACGACCGTG	4801
Db	4789	CAGAGACGGAAGCTTACATCTTACACCCCGGAGAGTCGCAAGTCCCAACACAGCATGTG	4848
OY	4802	CCCGAGACCAACCAACCCCATCTCGGCTTATGAGACCTGTCTTCCGAGGCGTGAAGTGGC	4861
Db	4849	CCCGAGACCAACCTCTCACCCCATCTCTCCCTTATGAGCACTTGTCTCCGAGGCGTGAAGTGGT	4908
OY	4862	GTCGACTGTATCGACGCAACATCCCTCGGCGTTTGACCCCACTTCATACCCCGCGGC	4921

[illegible]



Db 5932 GAGCCCGGATCCCTAGACACCCCGGAGCTCCAGCCACAGCCATCGCCCGACCCCGACA 5991  
QY 5999 AAGAACTTGGAACCTTACCAACGCGCAACCCCGCGGCGCACTGCTGGGCTCG 6058  
Db 5992 AAGAGCTTGGACCCCAACATGCGAGTCCGACCCCGCGG---CCCACTGGGCTTCA 6048  
QY 6059 GACCCGACCGGGAAAGACTCAAGATAAACCTTTTCCATCCAGGAACCTGAACTCCGT 6118  
Db 6049 GATCTGCATCGAAGAAAGACTCAAGATAAACCTTTTCCATCCAGGAATTTGAACTCCGT 6108  
QY 6119 TCTCTGGGTTAACCA---GGCAGACGTAACGCCCGGAAGGGGTGAGCCCGTCAAGCCCT 6175  
Db 6109 TCTCTGGGTTAACCAAGTGAAGCTGAGTCAAGCCCGGATGGGGTGGAGCCCATCAAGCCG 6168  
QY 6176 GTGAGCTCAACCCAGTGTGACCCAGCAAGAGGGGCTCCCAAGCACTGGAAAGGCTGAC 6235  
Db 6169 GTGAGCTCCCGGAGCTGACCCAGCAAGAGGGGCTCTCCAAACCTTGGAAAGGCTGAG 6228  
QY 6236 AAGAGCACTGGAGGGGGAGCTGCGGCGCCAGACAGCGGCGCGTGAAGCTTGGCGGG 6295  
Db 6229 AAGAGCACTGGAGGGGGAGCTGCGGCGCAAGACAGCGGCGCGCATGAAGCTCAAGCGG 6288  
QY 6296 GAGGCGGCGCACTCCACACCTGCGGCGCTGAGAGCCAGCCCTGTCAGCCCG 6355  
Db 6289 GAGGCTGCCCATCTCCACACTGCGGGCACTGCGCGAGGCGAGCCCTCATCAAGCCCA 6348  
QY 6356 CTGCTCAAGCGGCGCGGAGGCTCAAGGTCACAGCGGGGTGTCACTGCGCCAGCAC 6415  
Db 6349 CTCTCTCAAGCTGCGCGGAGCATCAAGGTCACAGAGGGGTGTCACTGCGCTCAGCAC 6408  
QY 6416 ATCAGTGAAGGTATCAACAGAGCTACACCGGCAACCCAGACAGAGCTCAGCGACCC 6475  
Db 6409 ATCAGTGAAGGTATCAAGAGGACTACACCGGCAACCCAGACAGAGCTCAGTGGCCCC 6468  
QY 6476 CTGCGCGCGCGCTCTACTCTCTCTGCGGGCGAGCTGCGCGCTGGAACCTCGCGCG 6535  
Db 6469 CTTCGCGCGCGCTCTACTCTCTCTTCCGAGACAGCTGCGCTGTGGATCTTCCGCGG 6528  
QY 6536 CCAACCAAGTGAACCTTCACTCTCCGCGCGGAGCAATGATGTCGCGCGCTGCTCCCC 6595  
Db 6529 CCAACCAAGTGAACCTTCACTCTCCGCGCGGAGCAATGATGTCGCGCGCTGCTCCCC 6588  
QY 6596 CACAGGAAGGGGGGCAAGAGGTCTCCAGAGCAAAAGAGGTCGAGCTTGGGGTGGT 6655  
Db 6589 CACAGTGAAGGGGGCAAGAGGTCTCCAGAACCCAGCAAAACATCGGCTCTGGCAGAGT 6648  
QY 6656 GAGGACGATTAAGAACCTGTGTCCCAACCGAGGGGATGACGAGGCGAGGCACTCCCG 6715  
Db 6649 GAGGATGGCATTTGAGGCTGTGTCCCAACCGAGGGGATGATGAGCCAGGACATGCTCG 6708  
QY 6716 AGTGTGTGTACCGCTGCTGTACCGGGATGGGAAACAGCGAGCCCAAGAGATGGG 6775  
Db 6709 AGGCTGTGTACCGCTGCTGTGTGAGAGCGGGAAACAGGGCGAGCC---CAGAGTGGG 6765  
QY 6776 TCCAACTTCCAGGCAACCAAGCCAGCGGCGGAGCTCTTCCAGCAAGCTGACCGAGG 6835  
Db 6766 TCTAACTTCCAGGCAACCAAGCCAGCGGCGGAGCTCTTCCAGTAACTGATCTGAGAG 6825  
QY 6836 AACTCCGCAATGATCAAGTCCAGAGAGCAGAGATCAACAAGAGCTGAAACCCCAAC 6895  
Db 6826 AACTCCGCAATGATGATGATGAGAGAGCAGAGATCAACAAGAGCTGAAACCCCAAC 6885  
QY 6896 CCGAATGAGCTGATCAATATATCAAGCAGCTGGGAGCGAGATCTTCAATATGCGCGC 6955  
Db 6886 CCGAATGAGCTGATCAATATATATGAGCAGCTGGGAGCGAGATCTTCAATATGCGCGC 6945  
QY 6956 ATCAACCGGAACAGGCTTATGACTATGAAAGCAGCGGCTGACAGAACTGACAGACC 7015  
Db 6946 ATCACTGAGAGCAGGCTTATGACTTATGAAAGCAGCGGCTGACAGAACTGACAGACC 7005  
QY 7016 AACATGGGCTGAGGCTCAATATTAGAAAGCACTCAGGGTAAATATGACAGTGGGA 7075  
Db 7006 AACATGGGCTGAGGCTCAATATTAGAAAGCACTCAGGGTAAATATGATCAGTGGGA 7065

QY 7076 GAGTCCCGCGCTGAGCGCAATGCTTTAAACCTCTGATGACAGTGCAGGCTGCCC 7135  
Db 7066 GAGCCCCCGCGCTGAGCGCAATGCTTTAAACCTCTGATGACAGGCTGCGCC 7125  
QY 7136 ---GCTGTATGGCCATTAACCGCTGTGACGAGCGAGTACCACTCACTGCGCA 7192  
Db 7126 GGTGCTGTATGCGCATTAACACTGTCTGACGAGAGTGAACCACTCACTGCGCA 7185  
QY 7193 GGTGCGGCGGGAAGGCTCAAGTCTTGGCAGACCCAGACCGGAAAGCAAGTCCCG 7252  
Db 7186 GGTGAGGTGGGAAGCCCAAGTCTTGGCAGACCTGAGAGCCGAAAGCAAGTCCCG 7245  
QY 7253 GCGCGGCGCTGACATGAGGAGACCGGCGCAACCTGTGTCTCTCAGTGCATCGAGGAG 7312  
Db 7246 GCAACAGGCTTACGCTGCGAGAGCGAACCCCTCTGTGTCTCTCACTGACAGAGGG 7305  
QY 7313 GACTGCAACCGCGGAGCGCGCTCAACCAACCGGTGTGGAAGACAGGCTTGTCCGA 7372  
Db 7306 GACTGCAATCGCGAATCACTCACTCAACCGGTGTGGAAGAGCGGCTTGTCACTGCA 7365  
QY 7373 GGTTCACGCTATTCCTTCAACCCCTGTATCATGCGCTGAGGCGGTGTATGGCT 7432  
Db 7366 GGTTCACGCTATTCCTTCAACCCCTGTATCATGCGCTGAGGCGGTGTATGGCT 7425  
QY 7433 TCCCGACCGCGCGCTCCCGCGGAGCGCGGCGCTGAGCGCGCGCTGAGCGCGCAAGCC 7492  
Db 7426 TCCCGCGCGCGCGCTGAGCGCGGAGCGCGGCGCTGAGCGCGCGCTGAGCGCGCAAGCC 7485  
QY 7493 TGGGACGAGAGGCGCAAGCCACTGTCTGTGCTGACATGAGACACTTCCGACAGAG 7552  
Db 7486 TGGGATGAGAGGCGCAAGCCACTGTGTGTTCAGATGAGACACTTCCGACAGAG 7545  
QY 7553 TGACTCAAGACAGGCGGCGGCGGCGGCGGTGEBERSFBLUTGATGCTCCAGGAGC 7612  
Db 7546 TGAACACGAGATGGGGGGAGGCGG-----TGCAGATCCCGACAA 7586  
QY 7613 CACAGGAAGGCGCTGACAGAGGCGGCGGTGCGGACTCCCGCAACCAAGAGAGGCC 7672  
Db 7587 GGCAGAGAGGCGCAACAGTGAAGACAGACAGCTGTGATCTCCGAGACTGAGAGAGGCC 7646  
QY 7673 CCTGATCGCGCTGC-----GCTCCATCATCTGTCCGTCCAGAGCGG 7717  
Db 7647 CCTGATCTGCTGCGCGCTGATCCGTCGTCGTCATCTGTGATCTGATCCAGAGCTGG 7706  
QY 7718 CATCTTGCCTGTCTAAAGCTTAAAGACTCCCGCGCGGCTGCGCTGTGAG-- 7775  
Db 7707 CATCC-TGCCTGTCTAAAGCTTAAAGACTCCCGCGCGGCTGCGCTGTGAGT 7765  
QY 7776 --ACCTTACTCAGGGAGTGTAACTGTGTGCTGCGGAGAGGGAG----- 7821  
Db 7766 ACCTTACTCAGGGAGTGTAACTGTGTGCTGAGAGGGAGTGAACAGAGGGAGGG 7825  
QY 7822 ----GGGCGGGGAGGGGAGCGGCAAGCGGTGTGAGCGGCAACAGGCGGCGGAGG 7876  
Db 7826 ACAGCGGGCGAGAGGGGGGGAGCAGCAATCGTGTGTGATGAGCTGATCTGGGG 7885  
QY 7877 CGGCAAGGACCCAAAGCAGATGACAGCACTTCAAGCGCACTGCT--CCCGGAAT 7934  
Db 7866 TCAGCGGGAGCCCAACAGAGCTGACCAAGCACTCAGTGCACCGCTGCGCTTAACT 7945  
QY 7935 GCATTTGGAACCAAACTCTAACTGAGCTGCAAGCCCGCGCGCTCCCTGCGGCTCCCA 7994  
Db 7946 GCATTTGGAACCAAACTCTAACTGAGCTGCAAGCTGCGGCTGCGCTTCTGTGCGCCCA 8005  
QY 7995 TCCCGCTTACGCTGTGACAGATGAGCGAGGCTGTGACGCGCCCAAGTGCCTGTT 8054  
Db 8006 GCGCGCTTGTGCTGTGACAGAGAGGTTCCAGCTTATCTGCTTAAATGCTGTATCA 8065  
QY 8055 CCGGTCGCCACAGACTGCGCCAGCCCAAGAGATGCTGAAACCAAGTCAAGGCGAGTGG 8114  
Db 8066 TCGCACTTCAAA---GGCAGCCAGGCCCAAGAGCTGAGGCGCATGACAGAGTGG 8122



[illegible]

```

RESULT 12
US-09-819-104A-6
? Sequence 6, Application US/09819104A
? Publication No. US20030027137A1
? GENERAL INFORMATION:
? APPLICANT: Chen, J. Don
? TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
? TITLE OF INVENTION: AND USES THEREFOR
? FILE REFERENCE: UMG-030
? CURRENT APPLICATION NUMBER: US/09/819,104A
? PRIOR FILING DATE: 2001-03-27
? PRIOR APPLICATION NUMBER: 60/193,138
? PRIOR FILING DATE: 2000-03-29
? NUMBER OF SEQ ID NOS: 6
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 6
?
? LENGTH: 7386
?
? TYPE: DNA
?
? ORGANISM: Mus musculus
?
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1) ..(7386)
?
? US-09-819-104A-6

```

Query Match	56.0%	Score 4797	DB 3	Length 7386
Best Local Similarity	80.3%	Pred. No. 0		
Matches 6118; Conservative	0	Mismatches 1200	Indels 301	Gaps 30

QY	2	ATGTCGGAGCTTCACACAGCTGTGGACACAGAGTGAGAGGCGCATGAGCCCGGTACACCG	61
Db	1	ATGTCAGGATTCACACAGCTGTGGACACAGCATGGCGGGCTGTGAGCCCGGTACCCA	60
QY	62	CCCCACAGCCCTTTCCTTACCCAGTGCATATGCGCCCGGAGACGACAGGAGCTGGGCTCTG	121
Db	61	CCCCATGGCATCTCTTACCCGGGTGAGTATACCCGGTCCACACAGGAGTATGGGCTGTGTT	120
QY	122	GAGTACCAAGACCAACTCCCGGACTATGCTCCACCTGTGTGCCGGCTCCATCATCCAG	181
Db	121	GAGTACCAACACACCCCGGTGATTACCTTCACACTGTGATCCCGGTTCCATCATCCAG	180

1

OY	182	CCCCAGCGGGGAGGCGCCCTCCCTGCTGTCTAGATTCCAGCCCGGGAATGAACGGTCCAG	241
Db	181	CCACGAGAGAGCGCGCCTCACTGCTGTCAAGATTCCAGCTGGAGTGAACGGTCTTAG	240
OY	242	GAGCTCCACCTGCGGCGCAGAGTCCCACTCATACCTGCGCCGAGCTGGGAGATCAGAGATG	301
Db	241	GAGCTCCACCTGCGCCCTGAGTCCCGACAGTTCCTGCTGTAGCTGGGCAACCCGACATA	300
OY	302	GAGTTCAATTGAAAGCAGCGCCCTTGCGCTAGAGCTGTGCTGACCCCTGCTGCGACCG	361
Db	301	GAATTCAACGAGAGCAGAGCGCCCGCTCGAGAGCTACACCGAATACCTCTCTGCGCCCA	360
OY	362	TCACCCCTGCTGGCCACCGGCGCAGCGTCCGGGATCTGAAAGCTTCAACAAAGACCTTAGC	421
Db	361	TCACCCCTGCTGGCCACTGGGCAAGCGGAGTGAGGTCTTGAAGACTTTAACAAAGACCTTAGC	420
OY	422	CTGACCGGCAAGCTGGAACCGGATGTCCTCCCGACGCCCCCGACACTGAACCTGAGCTG	481
Db	421	CTGGCAGGCAAGCTGGAGCTGTGTCACTCCAGTCCCGCGACGCTGACCTTAGAGCTA	480
OY	482	GAGCTGTGCGCGCACGCGCTGTCCAAAGGAGGAGCTGATCCAGAAC--TGGACCGCGTG	538
Db	481	GAGCTGTGCGCGCATCTGACTGTCCAAAGGAGGAGCTGATCCAGAACAGATTGACGCGCGTG	540
OY	539	GACCGAGAGATCAACATGCTAGACAGAGATCTCTAAGCTGGAAGAAACACAGACAGACAG	598
Db	541	GACCGTAGATCAACATGTGTAGACAGAGATCTTCAAGCTGAAGAAACACAGACAGACAG	600
OY	599	CTGGAGGAGAGGCTGCCAAGCCGCGCCGAGCTGTGAGAAAGCCGATGCACCGCCGCTATC	658
Db	601	TTGGAGAGAGAGGCGCGCCAGCCGCGCACCCGAGAGCCCTGTGTGCGCACCAACCTATA	660
OY	659	GAGTCGAAGCACGCGAGCCTGTGTGACATCTTACGACGAGAACCCGAGAAAGGCTGAA	718
Db	661	GAATCAAAAGCACCGAAGCCTGTGTCAAGTACTTACAGTAGAAGACCGAAGAAAGCCGAA	720
OY	719	GCTGCACATCGGATTTCTGGAAGGCTGGGGCCCCAGGTGTGAGCTGCGCTGTACAACAG	778
Db	721	GCCGCAACACCGGATCTCTGAAGGCGCTGGGGCCCCAGGTGTGAGCTGCTCTGTACAACAG	780
OY	779	CCCTCCGACACCGGCGAGTATCATGAGAAATCAAAATTAACCGAGCGCATCGGGAAGAG	838
Db	781	CCGTCTGACACAGCGCAGTATCATGAAACATCAAAATTAACCGAGCGCATCGGGAAGAG	840
OY	839	CTAATCTTTGTACTTCAAGAGGAGGATATCAGCTCGGAAACATGGAAGCAGAAATTCTGC	898
Db	841	CTGATCTTTGTACTTTAAGCGSAGGAACACCGCGCGCAAGCAAGTGGGAACAGCGCTTCTGC	900
OY	899	CAGCGCTATGACACAGCTCATGAGAGGCTTGGAAAAAAAGGTGAGCGCATCGAAAAACAC	958
Db	901	CAGCGCTATGACACAGCTCATGAGAGGCGTGGGAGAAAGTATGAGCGCATGAGAAACAT	960
OY	959	CCGCGCCGCGCGGCGCAAGAGAGGAAGGTGGCGAGTACTACGAAAGCAATTCCTCGAG	1018
Db	961	CCGCGAAGGAGAGGCGCAAGAGAGGAAGGTGAGGAGTACTACGAAACATTTCCCGAG	1020
OY	1019	ATCCGCAAGCAGCCGAGCTTCAAGAGCGCATGTGACAGACAGGAGTGTGGCGCAGT	1078
Db	1021	ATCCGCAAGCAGCCGGGAGCTTCAAGAGCGCATGTGACAGACAGGAGTGTGGCGCAGT	1080
OY	1081	GGGCTTTCATGTCGGCTGCCCGCAGACGAGACGAGGTTCTCAGAGATCATGATGGCTCT	1138
Db	1081	GGGCTTTCATGTCGGCTGCCCGCAGAGTAGCATGAGGTTTCTGAGATCATGATGGCTTG	1140
OY	1139	TCAGAGCAGGAAACTTGGAGAGCAGATGGCGCAGCTGCGCGTGAATCCGCGCATGCTG	1198
Db	1141	TCTAGAGCAGGAAACTTGGAGAAACAGATGGCGCAGCTGCGCGTGAATCC--GCCATGTTG	1198
OY	1199	TACGACGCTGACGAGCGCATGAGTCACTAAATGAACCGGCGCTTATGGCGGACCC	1258
Db	1199	TACGACCG--GACCGACGAGGAGTAAAGTTTATCAACATGAATGAATCTCATGATGAACCC	1257
OY	1259	ATGAAGGTGTACAAGACCGCAGAGTATGAAACATGTGAGTGAACGAGAAAGACGC	1318

Q

1259 ATGAAGGTGTACAAAGACCGCCAGGTCA TGACATGTGGAGTGAGCAGGAGAGGAGACC 1318

Db	1258	ATGAAAGGTCTTCAAGAGACCGTCAAGGTTTCAACATGTGAGCGAGCAGAGAGGGA	CAAC	1317
Qy	1319	TTCCGGGAGAAAGTTCAATGCAACATCCCAAGAACTTTGGCCCTGATCGATCATCTTCTGGAG		1378
Db	1318	TTCCGTGAGAAAGTTTATATGCAACACCTTAAGAACTTTGGCCGTGATTTGCCCTCATCTCGAG		1377
Qy	1379	AGGAAGACATGTGGCTGAGTGGCTCTCTATTAATACTGACTTAAGAAATGAGAACTAT		1438
Db	1378	AGAAAGACGGTCCGTGAGTGTCTCTTATTAATACTGACCAAGAAAGAAATGAATAATTAC		1437
Qy	1439	AAGAAGCCTGTGAGACCGGAGCTTATCCGGCGCGCGGCAAGAGCCAGCGAGCAACACAGC		1498
Db	1438	AAGAGCTTGTGAGGCGGAGCTTATCCGGCGCGCTGGCAAGACCCAGCGAGCAACAGC		1497
Qy	1499	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCCGCAGCAGCAGAGAGAG		1558
Db	1498	CAACACACGACGACGACGACGACG-----ATGGCAGGAGGAGCCAGAGAGAG		1545
Qy	1559	AAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG		1618
Db	1546	AAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG		1605
Qy	1619	AACGACAAAGAGAAACCTCCCTTAAGAGAGAGACACACCACTTCAAGGAGAGAAAGAC		1678
Db	1606	AACAGAGAGAGAGAACTCAGCAGAGAGAGACAGACGACCTTGTGGAGAGAAACGAT		1665
Qy	1679	GAGAGAGAGAGCTGTGGCTCCAAAGAGCCGCAAACTGCAACAGCCAGCCAGGAGAGAGCAA		1738
Db	1666	GAGAGAGAGAGCCGTGGCTCCAAAGAGCCGCAAACTGCAACAGCCAGCCAGCCGCAAA		1725
Qy	1739	GGCCGCATCAACCGCTCAATGGCTTAATGAGGCCAAACGAGAGAGCCATCAACCCCGAG		1798
Db	1726	GGCGGTATCAACGCGCTCCATGGCCCAAGAGGCCAAACATGAGAGAGACGACCAACCCAG		1785
Qy	1799	CAGAGCCCGAGCTGGCTCCATGAGAGCTGAATGAGAGTTTCCGTGAGACAGAGAGAGAA		1858
Db	1786	CAAAATTCAAGAGCTGGCTTCATGAGAGATGAACAGAGATTCTCCCTGACATGAAGAGAG		1845
Qy	1859	ATGGAACACGACCAAGAAAGGTCTCTGGAAACAGCGCCGCAACTGATCGGCATGGCCCGG		1918
Db	1846	ATGAGAGACAGCAAGAAAGAGCCCTCTTGAGACATGGAGAGAACTGTACAGCCATTGGCCGC		1905
Qy	1919	ATGTGTGGCTTCAAGACTGTGTGCGACAGTGAAGACTTCTTCACTTCAACTAACAGAGAG		1978
Db	1906	ATGTGTGGCTTCAAGACTGTGTGCCAGGTGAAGAACTTCTTCACTTCAACTAACAGAGAG		1965
Qy	1979	CAGAACTCCATGAGATTTTGACAGACGACCAAGCTGAAGATGAGAGAGAGAGAGAGCGC		2038
Db	1966	CAGAACTCGACGAATTCCTTCAACAGACCAAGCTTAAAGATGAGAGAGAGAGAGAGCGCT		2025
Qy	2039	CGGAGGAAGAGAGAGAGAGAGCGCCGCGCGCGCCGACGAGAGAGGTGCAATTCGCCGCCGTG		2098
Db	2026	CGGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG		2085
Qy	2099	GTGAGAGATGAGAGATGAGAGCGCTCGGGCTGAGCGGAAATGAGAGAGAGATGGTAGAG		2158
Db	2086	GCTAGAGACGAAGAGATGAGAGCATCAAGCGCCCAAGTGCMAATGAGAGAGAGCTGGCGAG		2145
Qy	2159	GAGGCTGAGACTTATCATGCTCTTGGGAAATGAGGTGCCAAG-----GGGAAATGCAATGGC		2215
Db	2146	GAGGCAAGAGCTTCAAGGCTCTTGGGAAATGAGGTGCCAAGAGTTGGGGAGTGCAGTGGC		2205
Qy	2216	CCAGCCACTGTGCAACACAGTCAAGACCGAGAGCAATCCCTCTCTCAACTGAAGGCC		2275
Db	2206	CCAGCTGTGTGCAACACAGCTCTGTGATCTGAGAGTGTCCATCCCGCGTTCAAGAGCC		2265
Qy	2276	GCCAAAGACACAGGAGCAAGATGGAGCCCAAGGCCCAAGCACTTGGGCGCGCGAGCGGCA		2335
Db	2266	ACGAAGAGAC-----TGGGCTTAAACCACTGGCACTGAAGCATTTGGCCGCTGGC		2316
Qy	2336	CCCCAGGCGCAACCAACCAACCAACGAGAGACATCCGAGGCGCCCAATTGAGCCACCCG		2395

D	2317	ACCCAGCACCTGTTCTCTCTCCAGAGAAACCCGGAACGAGCCCTGGAGACCTCCCA	2376
O	2396	GCTCTGAAGCACCGAGCCCTTACGCCCCACACAGACCCCCATGCGCTTGTGACCT	2455
D	2377	GTCCCTGATGCCA---GTGGCCCAACATCCCAAGACTTCCCATCACTGCGCACCC	2433
O	2456	CCTCTGTGGTCCCCAAGAGAGAAAGGAGAGACCGAGACGCGCCCAGTGAG	2515
D	2434	CCGGCTACTGTGGACAAGATGAAACAAGAACCCCGGCTGCTCCAGCTCCCAACAA	2493
O	2516	GAGGGGAGAGACGAAGACCCCGCGGCTAG-----GACGTGCA	2557
D	2494	GATCCCAAGAGACGAAGATCTGAGGCCGAGAGATCCATGTGGGAAAGCAGAGAGCC	2553
O	2558	GTGACACAGAGGAAGCCGAGAGACCCGTCAGAGCAAGTGCACGAGAAACCGAGAG	2617
D	2554	GAGGCTCTGTGAGAGCCCGGAGAGTGAAGAAGTGACCAAGAGAGAAACGAGGA	2613
O	2618	GGGCCC---GGCCAAAGGCGAAGAGCGCGAGGCCCTGAGGCGACGGCCGAGGGGGCTC	2674
D	2614	GAGCCTGAAACAAGCGCAAGGAGACAGAGGCCATTGAACTGTGCTGAAGCACACTT	2673
O	2675	AAGCAGAGAAAGAGAGGGCGGAGCGGACGGGCGACCACTGCGCAAGAGCTCGGCGCC	2734
D	2674	AAG---GTGAGGAGAGGCTGTGACAGAGCACTGTGACAAAGGTTCCAGCTCAGTGTCC	2730
O	2735	CCCCAGGACAGCGACTCCAGTCTTACCTGCAGTGCAGACGAGTGAATGAGGCGGAGGC	2794
D	2731	ACCCAGGACAGTGACTCAAGTGCACCTGCAGTGCAGTGAAGTGAAGTGAAGAACCCGAA	2790
O	2795	GGCGACAAAGACCGGCTGTCTGTCCCAAGGCCCAAGCTCTCACCCCGACTGGGACCCC	2855
D	2791	GGTACAAAGGCAAGCTGTCTGTCAACAAAGGCCAGCTCTCACCCTGGAGATGCC	2850
O	2855	CGGGCAATGCTCAACCCAGAAACCACTGGACCTGAAGCAAGCTGAAGCAGCGGCT	2914
D	2851	CGGGCAGATACCTGCCCCAGAAACCGCTGAGCTTGAAGCACTGAACAGCAGAGACCC	2910
O	2915	GCCATCCCCCCCATATCAGGTCAACCAAGTCCATGAGCCCCCGGAGAGACAGACTTCC	2974
D	2911	GCCATCCCCCTTACC---GTACCAAGGTCCATGAGCCCCCGGAGAGACAGATACCC	2967
O	2975	ACCAAGCAGTCTCCCCCAAGCCCCACCGCACCGCAAACTTGCAGCCGGAAGAGCGAGCC	3034
D	2968	CCAAAGCAGTTCCTCCCTGTGCTTCCACCGACGACACTTACAGCAAGAGGTGACCTG	3027
O	3035	CCTCAGGACGCTGGCAGCAGAGCCCCCGGGGGCAAGAGAGCCGGGACCCCCCGCGAC	3094
D	3028	TCTCAGCAGTCTGGAGAGAAATCCAGTGCAGTGCAGATGCCAGCCCAAGTGTCTTGCAG	3087
O	3095	AAGAGAGC-----CTTCCGACGCGAGGCCAGAAAGTGCCTG	3131
D	3088	AAAGAGCAGAGAAACCCGCAATTCTTCCGGCTTTCCTCACTGAGAGGCGCAAAAGCTACCGA	3147
O	3132	GGAACCCCCCTTGTGTGACATTCCGGGCTGTCCCTTCCCGGTGCGGCTGATCA	3191
D	3148	CTGAGCCCCCAGCTGTGATCAT-----CGGCTGCTTCCCATCTTCCACGGAGGTGATCA	3202
O	3192	AGGCTCTCCCGCATAGCCCCCGGACCCCTTCAAGCTTCTCTTACAGTTCACATGTGACCAAC	3251
D	3203	AGACTTCCACAGGGC-----TGACCTCTTCTCTCTTCAACACCCCGGTACCCGC	3256
O	3252	TGCCCTTGGGCTCATGACACTGCGGCGCGCTGTGCGCGGCGCCACCACTTCTCA	3311
D	3257	TGCTCTTGGGCTTCAAGATATGTGCCCGGCGGCTGTGCGCG---TCCCCCATCTCTA	3313
O	3312	ACCGGCTCCCTCATCTCTCTTGCACAGACCCCAAGCTCTTGAAGGCAAAATGAGTG	3371
D	3314	ACCCCCACCCCTCATCTCTCTGCAAGCAATCCCGGCTACTTGAAGGCGAGCTGGGTG	3373
O	3372	CCATCTCCCAAGAAATGTGGTCCAGTCCACAGTCCCGTACTCAGACATATGCCAAGGCC	3431
D	3374	CCATCTCCCAAGGATATCATGTCCAGCTTGTGTGCTCACTCAGAGATGCCAAG---C	3430

QY	3432	CGGGGGCCCTGTGATCCCATATGGGGCTGGCCCTCCGACATGGACCCCAAAAGCTTGGACCTT	3431
Db	3431	CCATGAGGCCCTCTCAACCATATGAGGCTGGCCCTTGGCCGTGAGCCCTTAAGAAAGCTG----	3483
QY	3492	TCAGCCGAGTGAACAGAGAGCAGCTGTCTCCACCGGGGCCAAGCTTGGGCCACCGGAGAGCC	3551
Db	3484	-----	3483
QY	3552	TGGGGGTGTCACAGCCCGAGAGGGCTCCGTGCTGAGAGGACAGAGCTTGGGCTTCAGTTCC	3611
Db	3484	-----GGGACAGCACT--GGCTCCGCCA	3504
QY	3612	CGGGCCGAGAGCATACCAAAAGCAATTCACAGCACACGGGTGTCCTCTGGACAGCCGCATCA	3671
Db	3505	CCAGTGGAGGATACCAAGGAGCCTCCCATGATCCGGGCTGAGACGG-----CCCA	3557
QY	3672	CATACCGGCTCATCAACCAAGGACAGCCAGCTGAGTCCGTGACAAAGGGCACCATCA	3731
Db	3558	GCTATAGAGGCTCTATCAACCAAC--GCAAGCCCGCAGAGCTCTCTACAAAGGATACATCA	3616
QY	3732	CCAGGATCATTGGCGAGAGACAGCCCGAGTGGCTTGGACCGCGGCCGGAGAGCAGCCTGC	3791
Db	3617	GCAAGATCTGTGGTGAAGGACAGCCCAAGTCCCTTGAACCGGGCACAGAGGACACCTTGC	3676
QY	3792	CCAAAGGGCCACGTATCTACGAGGCAAGAGGGGCACTCTTGTGCTATAGAGGTGGCA	3851
Db	3677	CCAAAGGGCCATGTCAATCTATAGAGGCAAGAAAGGCAAGTCTTATCTATAGAGGTGTA	3786
QY	3852	TGTCTGTGACCCAGTGTCTCCAAAGAGAGCGGCAGAGAGCAGTCTAGAACCCCCATGAGA	3911
Db	3737	TGTCCGTGTCAAGTGTCTCTAAAGAGAGTGAAGAGAGCAGTCCGGGCCCAACCATAGGA	3786
QY	3912	CGGCGCGCCCCAAGCGCATCTATGACATGATGAGAGGCGCGTGGGCAAGCCATCTTCT	3971
Db	3797	CTGCCGCCCTTAAACGCACTTATGACATGATGAGAGGCGCGTGTAGGCGAGACTGTCACT	3856
QY	3972	CAGCGAGCATGAAGGTCTCATGGGCGGTGCATCCGCGGAGAGACAGACCCCCAC	4031
Db	3857	CAGCGAGCATGAGAGGACTCATGGGCGCGCCATCTCTGAGACG-----ACAGCCCC	3910
QY	4032	ACCTCAAAAGACAGCACCACTCCGCGGGTCCATCAACAAGGGATCCCTCGGTCTACG	4091
Db	3911	ACCTCAAGAGCAGCATCATCATCCGAGGCTCATACGCAAGGGCATCCGAGGTCTATG	3970
QY	4092	TGAGGGCAGAGAGACTACTCGGTGGGAGGCCAGCTCTTAAAGCGGAGGGCACGC	4151
Db	3971	TGAGGGCGCAGAGAGACTACTTAAAGCGCGGAGGCCAAGCTCTTGAAGGAGAAAGGCACAC	4030
QY	4152	CTCGGCCCCCAACGGCCCTCAGGGGACCTGAACGAGGCGCTCAAGACGCGAGGCC-----	4204
Db	4031	CACCTTCCCCACCAACCACTCGGGACCTGACTTGAGACTTACAGGCCCGGCGCCCTGACCC	4090
QY	4205	--CTGGGCCCTCTGAAGCTGAAGCCGAGCCCATGAGGGCTGTGTGGCCACCGTGAAGAGG	4262
Db	4091	CTGTGGGGTCCCCTGAAGCTGAAGCCGACTCAAGAGGGTGTGTAGCACTGTGAAGAGG	4150
QY	4263	CGGGCCGCTCATCATATGAGATCCCGCGCGAGAGCTGCGGCAACGCTCGAGCTGCCCC	4322
Db	4151	CGGGCCGCTCATCATATGAGATCCCGAGAGGAGCTGCGCGGCAACTTGAAGCTACCCC	4210
QY	4323	TGGCGCCGCGCGCTCAAGAGGGGCTCATACGAGAGGGGACCCGCTGAAGTACGCA	4382
Db	4211	TGGACCAACGAGCTCTGAAGAGGGTTCATACCCAGAGGACCCCACTCAAGTACGACT	4270
QY	4383	CCGGCGGCTCCACCACTGGCTTCAAAAGCAGAGCTACGCTCTCCCTCATCGGACGCCCG	4442
Db	4271	CTGGGAGCACTCTCACTGGGACCAAGAAACAGAGTGTGGCTCATATCGGACGCCCG	4330
QY	4443	GCGGAGGTTTCCACCCGTCGACCCGCTGATGTGATGGCCGAGCGCCCGGCACTGGAAC	4502
Db	4331	GCGGCGCTTTCCTGCGCTTCGACCCGCTGGACATATATGGCTGACGCGCCGGGCACTGGAGC	4390

Qy	450	GTGCTGCTACGAGGAGACCTGAAAGCGCGGACAGGGACCGCAGACAGCTCGGGGGGCT	4562
Db	4391	GTGCTGCTATGAGAGAGTCTGAAGAGCGGTCAAGGACACAGATGTGTCAAGGGGCT	4450
Qy	4563	CCATTGCGGCGCGCGCCCGGTCAATTGTGCTGAGCTGGTAAAGCCGCGCAGAGCCCC	4622
Db	4451	CCATCAACAGTGGGGGTCAAGTGTGTGTGCTGAACTGGGCAAGCAGCAAGGCCAC	4510
Qy	4623	TGACTATGAGACCACGGGGACCCCTTTGCGGCGCACTTCCACAGGTTTCCGCTGA	4682
Db	4511	TGACTTACAAAGCCACGGGGGACCTTCAACAGTCACTGCAAGTGGCTCCCTGTGA	4570
Qy	4683	CCATTGGGGAGCCACGCGCGGCTGAGAGAGGGGAGCTTTGGTCCAGCAAGGCAATCC	4742
Db	4571	CCACGAGGAGGCCACGCGCCCTTCAAGAAAGCAGCTTCTATCAAGCAAGGCGTCC	4630
Qy	4743	AGACCGAAAGCTGACGTGAGCGCTGTGAGATGCAATGCCAATCCCGCAGCACGCTGC	4802
Db	4631	AGGACCGGAAGCTGACATTTACCCCGGAGATGCGCAATCCCAACAGCACTGTGC	4690
Qy	4803	CCGAGCACCACCCAACCCCAATCTGCGCTTATGAGCACTGTCTCGGGGCTGAGTGGC	4862
Db	4691	CCGAGCACCACTCAACCCTATCTCCCTATGAGCACTTGTCTCGGGGCTGACTGTG	4750
Qy	4863	TGGAAGTATGAGAGGCAATCCCCGCGCTTGAAGCCCAACCTCAATACCGCGGCA	4922
Db	4751	TGGAAGTATGAGAGGCTGACATCCATTGAGCTTGTGACCACTTCAATACCGAGGAA	4810
Qy	4923	TCCCTGTGA---CGCAGCGCTGCTACTACTGCGCGCAACCTGTGCCCCCAACCCA	4979
Db	4811	TCCCTGTGAAGAGAGAGCGCAGCTTACTGTGCGCGCACTTGGCCCCAGCCCA	4870
Qy	4980	CCTACCGGACCTGTACCACTTAACCTTATCCGCGGCTACCCGACCGCGGCGTGG	5039
Db	4871	CCTACCGGACCTGTACCACTTAACCTTATCCGCGGCTACCTGACCGCGGCGCTGG	4930
Qy	5040	AGAGCGGAGACCAATCAATGACTGACTATCACTCGGACAGATGACCAACAAACGG	5099
Db	4931	AGAGCGGAGACCAATCAATGACTGACTATCACTCGGACAGATGACCAACAAACGTG	4990
Qy	5100	CCACCGGACCGGAGCGAGCTGATATGCTGAGGGGCTCTCGCGCGCGAGTCTGCG	5159
Db	4991	CTCGCGGACCGGAGCGAGCTGACTGACTGAGGGGCTCTGACCGGAGAGTCTGCG	5050
Qy	5160	TGGCACTCAACTACGTGCGGGGTCCCGAGGCAATGACCTGTCCCAATGTCACACC	5219
Db	5051	TGGCCTCAATTAATTCGCTGGGCCCAAGGGGATTAATGACTGTGCCAATGTCACACC	5110
Qy	5220	TGCTGTGCTCGTGGCCCCGAGACCAAGGACCCCAAGCACCGGCATAGACCGCTTGGCT	5279
Db	5111	TGCTGTGCTCGTGGCCCCGAGACCAAGGACCCCTGCAACCGGCATAGACCGCTTGGCT	5170
Qy	5280	ACCTCCCGCACGGGCGCCGAGCCCTTCAAGAGCGGCAAGAGCTCCCACTTCCCGAG	5339
Db	5171	ACCTCCCGCACTGGGCGCCCGACCTTCAAGAGCGGCAAGAGTCTCACTGTGCCAG	5230
Qy	5340	GAGGTCCAAACAATTGACAAAACCAACACAGTCTCGTCCGAGCGGAGAGGAGACC	5399
Db	5231	GAGGCGCCCACTACCTAGTAAACCACTGCACTCTTATCGAGAGGG-----	5280
Qy	5400	GGGATGAGAGCGGAGCGGGATCGGGAGCGGAAAGTCAATCTCTCACTGTCACACGA	5459
Db	5281	-----GAAAGGGAAGTGAAGCGGGAACGAGCAAGTTCATCTCACTGTCACACTA	5332
Qy	5460	CGGTGAGAGACCAATCTGGAACCTGTGACAGAGCAAGAGCGGAGCGGAGCGGCA	5519
Db	5333	CAATGAGAGATGACCACTGGAACCTGTGAGAGCAAGAGCGGGGCTGGGGGCA	5392
Qy	5520	GCAAGCGCGGGGTGGGGGACAGACAGCCGCCCTTCCACTCCATGCCCCACAGC	5579
Db	5393	GCAAGCG-----CGCGGCTCCACACCAACAGC	5422
Qy	5580	ACTGCGCACTTCCCTGGAGCCAGAGTGCCTTCAGACAGACCAAGTGTCTTCA	5639



```

: PRIOR APPLICATION NUMBER: US 09/798,586
: PRIOR FILING DATE: 2001-03-02
: NUMBER OF SEQ ID NOS: 2059
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 650
: LENGTH: 7534
: TYPE: DNA
: ORGANISM: Mus musculus
: US-10-087-192-650

```

Query Match	53.0%;	Score 4543;	DB 5;	Length 7534;
Best Local Similarity	78.4%;	Pred. No. 0;		
Matches 6091;	Conservative	1;	Mismatches 1320;	Indels 357; Gaps 40;

[illegible]

QY	1785	GGCCATCACCCCCAGAGAGCGCCAGGCTGGCTCCATGGAGCTGAATTGAAGTTCTCG	1842
Db	889	GACAGCAACCCCAAGAAAGTTTCAAGCTGGCTTCCATGAGATGAACAAGAGTTCTCG	948
QY	1843	CTGGAACGAAGAAGAAATGGAAAACGCCAAGAAAGTCTCTTGAAAACAACGCCGCAACTG	1902
Db	949	CTGGACCTGAGGAAGAGATGGAGACGAAAGAAAGGCTCTTGAAACAATGGAGGAACCTG	1008
QY	1903	GTCGGCCATGCGCCCGATGTTGGGCTTCAGACTGTGTCCGATGTAAACAATTCTACTT	1962
Db	1009	GTCAGCCATTGCCCCGCTAGTGGGCTTCAAGACCGTGTCCAGATGTAAAGACTTCTACTT	1068
QY	1963	CAACTTACAAGAAAGAGGAGAAACCTCCATGAGCTTGGACACAGACAACAGCTGGAAGATGA	2022
Db	1069	CAACTTACAAGAAAGAGGAGAAACCTCCAGCAAACTCTTACAGACACAAGCTTAAAGATGA	1128
QY	2023	GAAAGAGAGGAACGCGCGAGAGAAAGAAAGAAACGCGCGCGCGCCAGCGAGAGGC	2082
Db	1129	GAAAGAGAGGAACGCTCGAGAGAAAGAAAGAAAGACCCGACGTGCGGCGAGCGAGAGAC	1188
QY	2083	TGCATTCCCGCCCGTGGTGGAGATGAGAGATGAGAGCGCTCGGGCGGTGAGCGGAATGA	2142
Db	1189	AGCCTTCCCACTGCGCTAGAGACGAAGAGATGAAAGCAATCAGGCGCAAGTGCATATGA	1248
QY	2143	GAGAGAGATGTTGAGAGAGCTGAAGCTTACATGCTCTGGGAATGAGGTGCGCAGA--	2200
Db	1249	GGAAGAGCTGGCGGAGGAGGACAGAGCTCACAGGCTCTGGGAATGAGGTTCCAGAGT	1308
QY	2201	-GGGGAATGCATGTGCTCCAGCACTGTCAACAACAGTCAAGACAACGAGAGCATTCCTTC	2259
Db	1309	TGGGAGATGTGAGTGGCCCAACTCTGTCAACAACAGTCTGAATATCAGAGGTGTCCATC	1368
QY	2260	TCCCTCACTGAGGGCGCCCAAGGACACAGGGCAGAAATGGGCCCCAAGCCCCCAACCTT	2319
Db	1369	CCGCGCTTCAGAAAGCCAGAAAGACAC-----TGGGCTTAAACCACTGGCACTGA	1419
QY	2320	GGGCGCGAAGGGGCAACCCCCAGGCCCCACCCCAACCAACGAGAGACATCCGGGCCCC	2379
Db	1420	AGCATTTCCGCGTGGCCACCAGGCACTGTCTCTCTCCAGAAAGAACCGGCAACACCCC	1479
QY	2380	CATTGAGCCCAACCCCGGCTCTGAAGCAACCGGAGGCCCTTACGCCCCCAACACACCCCC	2439
Db	1480	TGCTGAACTCTCCCACTCCCTTATGTACA--GTGGCCACCATTCCTCCAGAGCTTCTCCC	1536
QY	2440	ATCGCCTCTGCACCTCTCTCTGTGTGCTCCCAAGAGAGAAAGAGAGAGAGACCCGAGC	2499
Db	1537	ATCACCTGCGCGCAACCCCGGCTACTGTGACAAAGATGAACAAGAAAGCCCCGGCTGTCC	1596
QY	2500	AGCGCCCCCAGTGAGAGAGGGGAGAGAGACGAAGCCCCCGCGCTGAG-----	2548
Db	1597	AGCTCCCCCAACAGAAAGATGCCAAGAGGCGAAGGTCTGAGGCGGAGAGATCGATGTGG	1656
QY	2549	-----GAGCTGACAGTGAACAAGGAAAGGCGGAGAGGCCCCGTCAAGAGCGAGTGCAC	2601
Db	1657	AAAGCCAGAGAGAGCCCGAGGCTCTTGAGAGAGCCCCCGAGAGATGTAAAGATGACACAA	1716
QY	2602	GAGAGAAAGCCGAGAGAGGAGC--GGCCAAAGGCAAGAACGCGAGAGCGCTGAGGCGAC	2658
Db	1717	GAGAGAGAACGAGAGAGAGCTTGAAGACAAAGCCAAAGGGCAACAGAGCCATTGAAACTGT	1776
QY	2659	GCGCGAGGGGGCGCTCAAGGCAAGAAAGAAAGAGGGCGGAGCGGACAGGGCCACACTGC	2718
Db	1777	GTCGAGGACCACTTAAAG--GTGAGGAGGGCTGTGATGCAAGGCAAGCTGTGACAAAGG	1833
QY	2719	CAAGAGCTCGGGGCGCCCCCAGAGACAGCACTCAAGTCTACTGTGACAGTGCACAGAGT	2778
Db	1834	TTCCAAGTCAAGTGGCCACCAGGACAGTACTCAATGCGCAACTCTGACATGTCGATAGAGT	1893
QY	2779	GGAATGAGGCGGAGGGCGGCGACAAGAAACGGCTGTCTGTCCAAAGGCCAGCTCTCTCAC	2838
Db	1894	GGAAGAAACCCGAAAGAGGTGACAAGGGCAGGCTGTGTATCAACAAAGGCCAGGCTCTCTCAC	1953

2839 CCCGATGCGCAGCCCCGGGCGCAATGCTTCACCCCGAAGCAGCTGAGACCTGAGGAGCT 2898  
1954 CCGGCTGAGAGATCCCGGGCGCAGTACTGCGCCGAGAGCCGCTGAGCCTGAGAGAGCT 2013  
2899 GAAAGCAGCAGCGGCTGCCATTCGCCCATCGAGGTACCAAAAGTCCATGAGCCCCCGG 2958  
2014 GAAAGCAGCAGCGGCTGCCATTCGCCCATCGAGGTACCAAAAGTCCATGAGCCCCCGG 2070  
2959 GGAAGGACGAGCTCCCAAGCAGGAGCTCCCGAGCCCGCCAGCCCGGAAAACCTGCA 3018  
2071 GGAAGGACGAGTACCCCAAGCAGTCCCGTGTGCTTCACCAAGCAGACCTTACA 2130  
3019 GCGGAGAGCAGCAGCCCTCAGCAGCTGAGCAGCAGCCCGGGGCAAGAGCAGAGCC 3078  
2131 GCCAGAGGGTGAAGTGTCTCAGCAGTCCGAGAGAAATCCACGTGGCAAGTCCCGCAGCC 2190  
3079 GGCACCCCCCGCGCAGAGGA-----GACCTTGCAGCCGA 3114  
2191 AGTGCCTCCTGCGGAGAAAGGAGCAGAGAAACCGGATTCCTTCCGGCTTCCCACTGA 2250  
3115 GCGCCGAGAGCTGCGGCGGAGCCCGCTTGTGCTGAGACTTCGCGGCTGCCCTTCCCGGTC 3174  
2251 GGGCCCAAGCTACCGACTGAGCCCGCAGCTGATCGAGGCTGCTCCCTCCCAATCC 2310  
3175 CCGCCGTGAGGTGATCAAGAGCTCCCGGATGCGCGGACCCCTCAGCCTTCTCTAAGC 3234  
2311 TCCAGGGAGGTGATCAAGACTTCCCAACAGCGCGGTGACCCCTGTGCTTCTCTACAC 2370  
3235 TCCACTGTGTCACCCACTGCGCTCGGAGCTCCATGACACTGCGCGCGCTGTCGCG 3294  
2371 ACCCGCGGTCACCCGCTGCTCGGAGCTCCAGATGATGCGCGCGCTGTCGCAAG 2430  
3295 CCCACCCACATCTCCACACCCGCTCCCTCATCTCTCTGCGCAAGACCCCGAGCTCT 3354  
2431 ---TCCCCCATCTCTAACCCCCCAACCCCTCATCTCTGCGCAAGACTCCCGGCTACT 2487  
3355 CGAGAGGCAATATAGTGCATCTCCCAAGAAATGTGAGTCCAGCTCCAGCTCCGCTACT 3414  
2488 TGAAGAGCAGCTGAGGTGCATCTCCAGGGAGATGTAGTTCAGCTTCGATGCTTCACT 2547  
3415 AAGACATGCCAAGGCCCGGTGGGCGCTGTCAACATGAGGCTGCGCTGCCATGAGACC 3474  
2548 AAGACATGCCAAGGCCCGGTGGGCGCTGTCAACATGAGGCTGCGCTGCCATGAGACC 2607  
3475 CAAAAAGCTGCGACCTTCAAGCGGAGTGAAGACAGGACAGCTCTCCCAAGGGGCGAC 3534  
2608 TAAAGAGCTG----- 2617  
3535 TGGGCGACGAGAGGCTGGGGGTGCCACAGCCGAGAGGCGTCCGTGCTGAGAGGAC 3594  
2618 -----GGGAC 2622  
3595 AGCTTGGGCTCAGTTCGCGGCGGAGATCACCAAGGACTTCCGAGCAGCGGGTGC 3654  
2623 AGCAGTGGGCTCGGCCACGATGGAAGATCACCAAGGGGCTCCCGATACCGGG----- 2678  
3655 CTCGACAGCGGCATCATACCGCGGCTTCATCATCCCAAGGACGCGCAGCTACGTCCT 3714  
2679 --CTGAGACGGGCCGAGCTACAGAGCTTATCACCCAGCGCACCCCGCACAAGTCTCT 2736  
3715 GTACAAAGGACCATCATCAGAGATCATCGGCGAGGACAGCGCGAGTCCGTTGGACCGCG 3774  
2737 CTACAAAGGATCATATAGAGAGATCGTGGTAGAGCAGCCCAAGTCCGCTTGAACGGGCG 2796  
3775 CCGGAGGAGCAGCTTGCACCAAGGCGCAGTCACTTACGAGGACAGAGGCGCAGCTCTT 3834  
2797 ACGAGAGGACACCTTGCACCAAGGCGCATGTCACTTGAAGGCGAAGAAAGCGCACGTCCT 2856  
3835 GTCTCTATGAGGGGTGAGATGTCTGTGACCCAGTGTCTCAAGAGAGAGCGCAAGAGCTTC 3894  
2857 ATCTCTATGAGGGGTGATGTCTGTGACAGTGTCTTAAAGAGAGATGAGAGGAGCTTC 2916  
3895 AAGAGCCCCCATGAGAGGCGCGCCCAAGGCGACCTATGACATGATGAGAGGCGCGCT 3954

2917 GGGCCCAACCCCATGAGACTGCGGCCCTTAAACGCACTATGACATGATGAGAGGCGGTGT 2976  
3955 GGGCAGAGCATTCTCTTACAGCAGATTCGAAGTCTCATGGGCGGTGCCATTCGCCCGGA 4014  
2977 AGGCAAGACTGTTCACCTTACAGCAGATTAAGAGGACTCATGGGCGGTGCCATTCCTTGA 3036  
4015 GGCACAGAGCCCGCACCACTTCAAAAGAGCAGCAGCATTCGCGGGGTCCATTCACAGAG 4074  
3037 GC-----ACAGCCCGCACCTTACAGAGAGCAGATTCACATCCAGAGCTCATTCACGAG 3090  
4075 GATCCCTGAGTCTCATGTTGAGAGCAGAGAGCTTACCTGCTGAGAGGCGCAAGCTCTT 4134  
3091 CATCCGAGGTCTTATGTGAGAGCGCAGAGGACTTACTTACGCGGAGAGCGCAAGCTCTT 3150  
4135 AAAGCGGAGGCGCAGCGCTTCGCCCGCCACCGGCTTACCGGACCTTACCGAGGCTTACAA 4194  
3151 GAAAGCAGAAAGGAGCACACACCCCGCACCACTTCGAGACTGACCTGAGACTTACAA 3210  
4195 GACGCAAGCC-----CTGGGCGCCCTGAAGCTGAAGCGGCGCATGAGGGCGCTGT 4245  
3211 GCGCGGCGCCCTGAGACCTCTGAGGTCCCTTGAAGCTGAAGCGGACTTACAGAGGTTGTGT 3270  
4246 GGCACAGTGAAGAGAGCGGCGGCTCCATTCATGAGATCCGCGGAGAGGCTGCGGCA 4305  
3271 AGCACTGTAAGAGAGCGGCGGCTCTATCCATGAGATCCGAGAGAGAGGCTGCGCG 3330  
4306 CAGGCGGAGCTGCCCTTGGCGCCCGCGGCGCTCAAGAGAGGCTTCATCATCGAGAGGCG 4365  
3331 CACACTGAGCTACCCCTGAGCACACGCGCTCTGAAGAGGGTTCATCACCCAGGCGAC 3390  
4366 CCGGCTCAAGTGAACACGCGCGGCTCAACACTGACTCCAAAGACAGACTAGCGCTC 4425  
3391 CCACTCAAGTACACTTGTGGGACCTTCATGACACAGAAACAGACTGTGGCTC 3450  
4426 CTTATCGGCAAGCCCGGCGGAGCTTCCACCCGTCACCCGCTGAGATGTGATGGCGA 4485  
3451 CATCATCGGCAAGCCCGGCGGCTTCCGCGCTGCAACCGGCTGACATTAATGGCTGA 3510  
4486 CCGCGGCGCATGGAACGTGCTGCTACAGAGAGCTTGAAGAGCGCGGCGCAGGAGCCG 4545  
3511 TGCCTGGGCACTGGAGCTGCTGCTATGAAGAGTCTGAAGAGCGGCTCAGGAGCACG 3570  
4546 CAGAGCTGGGGGGGTTCATTGGCGGCGCGCGCTCATTTGAGCTGAGCTGGGTAA 4605  
3571 CAGTGTGAGAGGGGTTCATACAGTGGGCTTCAAGTGTGCTGATCTGAACTGGGCA 3630  
4606 GCGCGGCGAGAGCCCGCTGACCTATGAGAGCAGCGGCGCACCTTTCGCGGCCACTGCC 4665  
3631 GCGAGGCGAAGGCCCATGACTTACGAAAGACACGCGGCGACCTTACCAAGTCACTGCC 3690  
4666 ACGAGTTTCGCCCGTGAACAATGCGGAGCCCAAGCGCGCTGACAGAGGCGAGCTTTC 4725  
3691 ACGTGGCTCCCTGTGACCAAGAGGAGCCCAAGCGCATTCCTCCCTATGAGACTTGGCT 3750  
4726 GTTCAGCAAGGAGATCCAGAGCCGAAGGCTGATGAGCGCTCGTAGATGCGCAAGTC 4785  
3751 ATTCAGCAAGGAGTCCAGAGCCGAAGGCTGATTTAACCCCGGAGATGCGCAAGTC 3810  
4786 CCGCAGACAGCCTGACCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4845  
3811 CCGCAGACAGCCTGACCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3870  
4846 TGGGGCGTGAAGTGGCGTGAAGCTGTATGACAGCAGCATTCCTGCTTGAACCCGAC 4905  
3871 CCGGGCGTGAAGTGGCGTGAAGCTGTATGACAGCAGCATTCCTGCTTGAACCCGAC 3930  
4906 CTCATACCCCGGAGCATCTCTTGA-----CGAAGCGGTGACTTACTGCGCGGCA 4962  
3931 CTCATACCCCGAGAGATCTCTTGAAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3990  
4963 CTTGGCGCCCAACCCCATCTACCGGACTGTACCCAGCTTACCTGATCGCGGCTAACCC 5022



Db 3991 CTTGGCCCCAGCCCCCACTTACCACACTGTATCCCACTTACTCTCATCCGCGCTACCC 4050  
Qy 5023 CGACACGGGGGCGCTGAGAACCGGAGACCATCATCTAATGAATGACTATCATCTCGAGCA 5082  
Db 4051 TGAACAGGGGGCGCTGAGAAACCGCAGACCATCATCTAATGAATGACTATCATCTCGAGCA 4110  
Qy 5083 CATGACCAACAACAGGCCACCGGCATGGGCCAGGAGCTGATATGCTGAGGGGCTCTC 5142  
Db 4111 GATGACCAACAACAGCTGCTCGGCATGGGCCAGGAGCTGATGACTGAGGGGCTCTGTC 4170  
Qy 5143 GCCCCGAGTCTGTGCTGGCACTAAGCTGAGGGTCCCGAGGATCATGAGACT 5202  
Db 4171 ACCGGAGAGTCTGTGCTGGCCCTCAATTATGCGCTGGCCCAAGAGGATATGAGCT 4230  
Qy 5203 GTCCCAAGTGCCACACTGECTGTGTCTGTGCCCCGAGACCAAGGACCCCAAGCACGCG 5262  
Db 4231 GTCCCAAGTGCCACACTGCGCGGTGTGCAACAAGCAAGGACCCCTCGCACGCG 4290  
Qy 5263 CATGACCGCTTGCTTACTCTCCCAACGCGGCCCAAGCCTTACAGAGCTCGCAAGCA 5322  
Db 4291 CATGACCGCTTGCTTACTCTCCCACTGCGGCCCAAGCCTTACAGAGCTCGCAAGTAG 4350  
Qy 5323 CTCCCACTCTCTCCAGAGGATCAACACTTGAACAAACCAACAACAAGTCCGTC 5382  
Db 4351 CTCACTGCTGTCCAGAGGCTCCACTTACCTTAAACCAACTGCCACATCTTCATC 4410  
Qy 5383 CGAGCGGAGCGAGACCGGAGATGAGAGCGGAGACCGGATCGGAGCGGAAAGTCAAT 5442  
Db 4411 GAGAGCG-----GAAACGGGAAGTGAAGCGGGAACGAGAACATGTCAT 4452  
Qy 5443 CTCACTGTACACATGACGCTGGAAGACGCAACCATCTGAGAGACTTGATCAAGACAG 5502  
Db 4453 CTCACTGTACACATGACGCTGGAAGACGCAACCATCTGAGAGACTTGATCAAGACAG 4512  
Qy 5503 CAGCGGAGACGAGCGGAGCGAGCGGGGATGAGGGGAGACAGCAACCGCGCCCTCCCA 5562  
Db 4513 CAGCGGAGCGGAGCGAGCGG-----CCCGCG 4542  
Qy 5563 CTCCCATGCCACACGACTCGCCCATCTCCCTCGAGACCCAGAGTGCCTTCAGACAG 5622  
Db 4543 CTCCCAACCCACGACACTGGCCATCTCCCGGAGCCAGAGAGCTTGACACAG 4602  
Qy 5623 ACCGAGTGTTCACACACAGGATGAAGGATCATCACCGCTGTGAGCCCAAGCA 5681  
Db 4603 GCCAGTGTGTGCAACAACAAGCATGAAGGCGTGTGCTCACTCCGCTGGAACCCGCA 4662  
Qy 5682 AGCCACGAGTCTGAGTGCACCTGCACTCTCAACCGTTCGCGCAGTGCACATTC 5741  
Db 4663 CGCCACGAGTCTGAGTGCACCTCACTCTTTCGCTGTCCGCGCAGTGCACATTC 4722  
Qy 5742 CACTGCAACCACTGCACTGGCGGCACTCGATGAGGATCACTCACTCATGG 5801  
Db 4723 CACTGCAACCACTGCACTGGCGGCACTCGATGAGGATCACTCACTCATGG 4782  
Qy 5802 AGCCGTCTTGTGCTGCAAGAGAGCGCGGATGCGCGGCAAGAGCGCGCGAGCA 5861  
Db 4783 AGCCGTCTTGTGCTGCAAGAGAGCTGTGCGGATGCGCGGCGCGAGCGCGCTGTGG 4842  
Qy 5862 ACAACGAGCATGCTCTCTGCAAGCGCCCAAGCGCGCTCGCGGATGAGAGCGCGCTCT 5921  
Db 4843 ACGCTGAGCATGCTCTCTCAACAAACCCCGGCG-----GAGCGCGCTCTCT 4893  
Qy 5922 CCCCCAGAGAGTCTGAGGCGCGCGCTTATGATGCTCTGTCTGTGCAAGCAACA 5981  
Db 4894 CACCGAGAGAGTCTCGAGCGCGCAATCCATGACACCCCGAGCTCAAGCAACAAGCA 4953  
Qy 5982 TCGCCCGCACTCTGCAAGAACTTGCACCTGACACGCGAGCGCGGAGCGCGCGGCG 6041  
Db 4954 TCGCCCGCACTCTGCAAGAAAGCTTGCACCCCAATGCAAGTCCGAGCGCGCGGCG- 5012  
Qy 6042 CACGCTCTGCGCTCGAGCCCGCAACCGGAAACATCAAGTAAACCTTTTTCATCC 6101  
Db 5013 --CCACCTCGGCTCAAGTCTGACCGAGAAAGATCAAGTAAACCTTTTTCATCC 5070

Qy 6102 AGGAATGAACTCCGTTCTGTGGTTACAC---GGAGAGCTTACAGCCCGGAGGG 6158  
Db 5071 AGGAATGAACTCCGTTCTGTGGTTACACAGTGAAGTGGCTTACAGCCCGGATGGGG 5130  
Qy 6159 TGAAGCCGTCAGCCCTGTGAGCTACCGATGTATACCAAGAAAGGAGCTTCCCAAG 6218  
Db 5131 TGAAGCCATCAGCCCGGTGAGCTCCCGAGCTTACCAACAAGGAGCTTCCCAAC 5190  
Qy 6219 ACCTGAAGAGCTCGACAAGAGCCACTGAGAGGGGAGTGGCGGCCAAGACCGAGGCC 6278  
Db 5191 CTCTGAAGAGCTTGAAGAAAGCCACTTGAAGGGAGTGGCGGCAACAGCCAGGCC 5250  
Qy 6279 CCGTGAAGCTTGGCGGAGAGCGCGCCCACTTCCCAACCTGCGGCGCTGTCTGAAGCC 6338  
Db 5251 CCATGAAGCTCAGCGCGGAGGCTGCCATCTCCCACTGTGGGCGCACTGCGCGAGAGCC 5310  
Qy 6339 AGCCCTCGTCCAGCCCGCTGTCCAGACCGGCCCAAGGATCAACAGCGCGGTGG 6398  
Db 5311 AGCCCTCATCAGCCCACTCTCCAGACTGCCCCAGGATCAAAAGTCAACAGAGGTTGG 5370  
Qy 6399 TCACCTGAGCCAGCATCATGAGGTCAATCAACAAGACTACACCCGAGCAACCCAC 6458  
Db 5371 TCACCTGAGCTCAGCATCATGAGGTCAATTAAGAGACTACACCGGCCACCCGC 5430  
Qy 6459 AGCAGCTCAGGCACTCCGCGCGCCCTCTACTCTCTTCCCTGAGGCGAGCTGCCG 6518  
Db 5431 AGCAGCTCAGTGGCGCCCTTCCCGCCCTCTCTACTCTCTTCCGAGCGAGCTGCCG 5490  
Qy 6519 TCTGGAACCTCGCGGCCCAACCATGATCACTTACTCTCCGCCCCCGGACCATGATGCC 6578  
Db 5491 TCTGGAATCTTCGCGGCCCAACCATGATCACTTACTCTCCACCCCGCACATGACAC 5550  
Qy 6579 CGGCGCGTGGCTCCCGCACAGCGGAAGGGGCAAGAGTTCAGAGGCAACAAGAGT 6638  
Db 5551 CAGCCCGGAGATCCCGCACAGTGAAGGGGCAAAAGTCCCAAGACCCAGCAAAAT 5610  
Qy 6639 CGGTCTTGGATGTGTGAGACGGTATTTGAACCTGTGTCCCAAGGAGGATGAGCG 6698  
Db 5611 CGGTCTTGGAGGAGGATGAGGATGACATTTGAGCGTGTTCACCAAGAGGCAATGACG 5670  
Qy 6689 AGCCAGGCACTCCCGGAGTGTGTATCCGCTGCTGTACCGGATGGGGAACAGACGG 6758  
Db 5671 AGCCAGGACATGCTCGAGGCGCTGTGTACCACTGTATTCAGAACGGGGAACAGGCG 5730  
Qy 6759 AGCCAGGAGATGGGCTCAAGTCTCCAGGCAACACAGCAAGCGCGCAAGCTTCTCA 6818  
Db 5731 AGCC---CAGATGGGCTTAAATCTTCCAGGCAACAGCGCGCGCAAGCTTCTCA 5787  
Qy 6819 GCAAGCTGACCGAGAGCACTCCGCAATGTCAAATCCAAAGCAAGAGATCAACAAG 6878  
Db 5788 GTAAGCTGACGAGAGCAACTCCGCAATGATGAAGTGAAGAGAGAGATCAACAAG 5847  
Qy 6879 AGCTGAACCAACAACCGGAATGACTGAATACATATCAAGCAAGCTGTGAGCGAG 6938  
Db 5848 AACTCAACCAACAACCGGAACGAGCCAGATACAAATATGAGCGCTGTGAGCGGAA 5907  
Qy 6939 TCTTCAATATGCGCGCATCAACCGGAACAGGCTTATGACCTTATGAAGCAAGGCGGCG 6998  
Db 5908 TCTTCAATATGCGCGCATCACTGAGAGGAGCTTATGACCTTATGAAGCAAGGCGGCG 5967  
Qy 6999 AGGAATGACCAAGCAACAATGAGGCTGAGAGCCATATTAAGAAAGCACTCATGAGTA 7058  
Db 5968 AAGAAACAGCGACAGCAACAATGGGGCTGAGAGGCAATTAAGAAAGCACTCATGAGTA 6027  
Qy 7059 AATATGACAGTGGAGAGTCCCGCGCTCAGCGCAATGCTTTTAACTCTGAATG 7118  
Db 6028 AATATGATCAATGGAGAGAGCGCGCGCTCGAGCGCAATGCTTTTAACTCTGAATG 6087  
Qy 7119 CCAATGCAAGCTGCC---GCTGCTATGCGCATTAACCGCTGTGAACGAGAGTGAAC 7175  
Db 6088 CCAAGCGCAATGCTGCCGCTGTCTATGCCCATTAACCACTGTGAACGAGAGTGAAC 6147

QY	7176	ACACATCACTCCGCGCAGAGTGGCGGGGAGAGGCGCAAGTCTCTGGAGACCCAGCAAGCC	7235
Db	6148	ACGCCTCACTCCGCGAGGTGAAGTGGAAAGCCAAAGTCTCTGGAGACATAGCAAGCC	6207
QY	7236	GAAAAGCCAAATCCCGGGCCCGGGGCTGGCATCTGGGAGCCGGGCAACCCCTCTGTCTCT	7295
Db	6208	GAAAAGCCAAATCCGCGAGCAACAGGCGCTAAGGTCTCGAGAGCCGACCCCTTCTGTCTCT	6267
QY	7296	CAGTGCATCTCGAGGGAGACTTGCACAACGCGCGGAGCGCGCTCAACCAACCGCTGTGGAGG	7355
Db	6268	CAGTACACTCAGAGGGGGAGCTGCATATCGCCGAACACCACTACCAACCGTGTGGAGAGG	6327
QY	7356	ACAGGCGCTCGTCCGCGAGGTTCACAGCATTCCTCCCTACAAACCCCTGATCATATGGGGCTGC	7415
Db	6328	ACCGGCGCTCATCTGCAGAGGTTCACAGCATTCCTCCCTACAAACCTTGTATATGAGGCTAC	6387
QY	7416	AGGCGGGGTGATGAGCTTCCCAACCCCAACGGGCGCTCCCGCGGGAGCAGGGGCGCTCG	7475
Db	6388	AGGAGAGGTGTATATGAGCTTCCCGCGCGCCCAACCTTGGGCTTGGCGAGCAGCGGGGCGCTT	6447
QY	7476	CTGGGCGCCCAACCAAGCTGGGAGCAGAGAGCCCAAGCCACTGCTCTGCTCGCATACAGAG	7535
Db	6448	CTGGTCCCCCAACCAAGCTGGGATGAGAGGCCAAGCCACTGCTGTGTTCAACAGTATGAG	6507
QY	7536	CACCTCTCGAGACAGGAGTGAATCTAGAAACAGGGCGGGGGGGGGCGGCGGTUEUS.PEPLU	7595
Db	6508	CACCTCTCGAGACAGGAGTGAACACAGGATATGGGGGGGAGCGG-----T	6549
QY	7596	GTCAGGTTCCAGCGAGGCCACAGAAACGGGCGCTGAGAGAGCGGGGCGGTGCGACTGCC	7655
Db	6550	GCCAGGTCCCGCACAAGGACAGAAAGCAGCCGAGATGAGCAGACAGTGTGACTGCC	6609
QY	7656	CAACCAAGGAAGAGAGCCCTGTGAGTCCGCTGCGCTCCATC-----CAT	7699
Db	6610	AGACTGAGGAAGAGAGCCCTGTGAGTCTGGGTGCATCCGTCGTCGCTCCATCAT	6669
QY	7700	CTGTTCCTGTCAGAGCGCGGATCTCTTGCTGTCTAAAGCTTAACTAAAGCTCCGCCCC	7759
Db	6670	CTGTTCATCTCAGACTGGCATTC-----TGCTGTGTCTAAAGCTTAAAGCTCCCAACCCG	6728
QY	7760	GGCTGGGCGCTGTGAG-----ACCTTACTCAGAGGAGTGTTAACCTGGAGCTTCGGGAAGGA	7815
Db	6729	GGGCTGGGCGCTGTGAGCCTTAACCTCAGAGGAGTGTTAACCTGGAGCTTCGAGAGAGGA	6788
QY	7816	GGGGA-----AGGGGCGGGGAGGGGGGCACGCGCAGGCGGTGTGGCAGCC	7858
Db	6789	GTGGAACAGAGAGGAGGAGCAGCGGCGCAAGAGGGGGGGGAGACAGCAATCGTGTCTAGTC	6848
QY	7859	ACACACAGAGCGGCGCAGAGCGCGCAGAGGACCCAAAGCAGAGTACACGCACTTCACGCG	7918
Db	6849	GCACTCGGTGATCTGGGGCAG-----CGGGGACCCACACAGGCTGACCAAGCACTTCATGCGC	6907
QY	7919	ACTGCGCT-----CCCCGGAATGCATTTTGGAAACCAAATCTTAACTGAGCTTCGACCCCGCG	7976
Db	6908	ACGCGCTCGCCCCCTTACCGCATTTTGGAAACCAAATCTTAACTGAACTCTCGCGTGTCT	6967
QY	7977	CCCTTCCTCCGCGCTCCCATCCCGGCTTAGCGCTCTGCAGACAGATGAGCGAGGCGCTGTGCA	8036
Db	6968	GCGCTCTCTTGTGCGCGCCCAAGCGCGCTTGCGCTCTGCAGACAGACGTTCCAGCTTATC	7027
QY	8037	GCCCCCAAGTGCCTGTTCCGGGTCCCAACAGACTGCCCAAGCCCAACAGATTTGCTGAGAA	8096
Db	7028	CTGCCCTTAATGCTGTCAATCGCAGTCTCCAAAGGCCAACCCAGCCCAAGACTGGAGAG	7087
QY	8097	CCAAAGTCAAGGCTCAGGTGGCGGAGCAAAAAGGCGCAGGTGCGGCTTGGGGGGGAAAGGATGCT	8156
Db	7088	CCCA--TCAGACAGAGTGGGTGCAC--AAAGGCGCAGGCGCTGGCTGG--GGACAGGATGCT	7143
QY	8157	CCGAGGACTGTGACTGTTTTTTTTCACACATCGTTGCGCGAGGCTGGGGAAGGAAGGAGAGA	8216
Db	7144	TGCAGGAAGTGGACCGTTTTCCCGGCGCTGTGTCTGTGGCAC-----GGAGGGAAGAGCAG	7199
QY	8217	TGTAAATGATGTGTGGTTTTACAGGAGTATATTTTTGATACCTTCAATGATTAATTGAGA	8276

Db	7200	TGTTAAAG--GTGTGGCTTACAGGGATATTTTGTATACCTTCAATGAGTATATTGAA	7257
Qy	8277	TGTTTACGCAAGAGAGAACTTACCAGATATTACTGTCTGCTGTTGATCTTGCTT	8336
Db	7258	CGTCTCAACAAGAGAGAACTCGCC-GTGTATTTGTGCGCTGTGTC-TTGTGCTCTACT	7315
Qy	8337	ACCGTTCAAGAGAGCGTGTGAGAGCGGCACATGTCGGTGAACCCATATCACTCGACAGCAAG	8396
Db	7316	ACTGTTTCAGAGGACCTGCAAGC--AAGCGGTGGCCCAACCATACGACAGACTTGGG	7373
Qy	8397	GGGCGGGAGACTGCTCTGTACAGCCGCCGCTG-TGTCTCTCCCTCCCTCCCTCTTGGGAGA	8455
Db	7374	GGTAAAGGGGCTCCGGGACACAGGCACTGTGCCCTTCCCCCAACCCCTTCCCTTGGGAAA	7433
Qy	8456	ATGAATTGATGGGTATTTGTGTGGCCGCCATTGGCGAGAGGGTGTGTATTTCTGTATT	8515
Db	7434	ATGGACTTGATGGGTATTTGTGTGGCCGCCCTCTGTGTGACGG---CGGCATTGTGTATT	7490
Qy	8516	ACAACAAGTGGTCTTAATTAAGAGCAAT	8544
Db	7491	ACAATGTGTGTCCAAATTAAGAGCAAT	7519

```

RESULT 14
US-10-146-473-14
Sequence 14, Application US/10146473
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Scanlan, Matthew
APPLICANT: Gout, Ivan
APPLICANT: Stockert, Elisabeth
APPLICANT: Gure, Ali
APPLICANT: Chen, Yao-Tseng
APPLICANT: Old, Lloyd
TITLE OF INVENTION: Breast Cancer Antigens
FILE REFERENCE: L00461/70130(JRV)
CURRENT APPLICATION NUMBER: US/10/146, 473
CURRENT FILING DATE: 2002-05-15
PRIOR APPLICATION NUMBER: US 60/291,150
PRIOR FILING DATE: 2001-05-15
NUMBER OF SEQ ID NOS: 82
SOFTWARE: PatentIn version 3.0
SEQ ID NO 14
LENGTH: 2930
TYPE: DNA
ORGANISM: Homo sapiens
US-10-146-473-14

```

Query Match 30.6%; Score 2618.6; DB 5; Length 2930;

BEST LOCAL ALIGNMENT 24.08; E-VALUE NO. 0  
 Matches 2786; Conservative 3; Mismatches 9; Indels 146; Gaps 2,

Qy	4785	CCCCGCACACCGTGTCCCGAGACACCCACACCCCATCTGGCCTTAGAGACCTGC	4844
Db	133	CCCCGCACACCGTGTCCCGAGACACCCACACCCCATCTGGCCTTAGAGACCTGC	192
Qy	4845	TTCCGGGCGGTGAGTGGCGTGTGACCTGTATGTGCAGCCACATCCCCCTTGGCCTTGAACCCCA	4904
Db	193	TTCCGGGCGGTGAGTGGCGTGTGACCTGTATGTGCAGCCACATCCCCCTTGGCCTTGAACCCCA	252
Qy	4905	CTTCACATACCCCGGCGCATCCTCTTGAGCGACGCGGTGCCTTACTACTCGCCCGACACC	4964
Db	253	CTTCACATACCCCGGCGCATCCTCTTGAGCGACGCGGTGCCTTACTACTCGCCCGACACC	312
Qy	4965	TGGCCCCCAACCCGACCTTACCCGACCTGTAAACGACCTTATCCGCGGCTAACCCCG	5024
Db	313	TGGCCCCCAACCCGACCTTACCCGACCTGTAAACGACCTTATCCGCGGCTAACCCCG	372
Qy	5025	ACACGGGGGGGTGAGAAACCGGACAGCACTATCATATGATCTATCATACCTCCGAGAGA	5084
Db	373	ACACGGGGGGGTGAGAAACCGGACAGCACTATCATATGATCTATCATACCTCCGAGAGA	432

QY 5085 TGCAACAACAACGAGCCACCCGCAATGCGCCCAAGAGTGTATATGCTGAGAGGCTCTCGC 5144  
| | | | |  
Db 433 TGCAACAACAACGAGCCACCCGCAATGCGCCCAAGAGTGTATATGCTGAGAGGCTCTCGC 492  
| | | | |  
QY 5145 CCCGGAATCCTCGCTGGCACTCAACTAAGCTGCGGGTCCCGAAGCAATCACTGACTGT 5204  
| | | | |  
Db 493 CCCGGAATCCTCGCTGGCACTCAACTAAGCTGCGGGTCCCGAAGCAATCACTGACTGT 552  
| | | | |  
QY 5205 CCAAGTGCACAACATGCTGTGTCTGTGTCCCGCCCAACAAGCAACCCAGCCACCGCA 5264  
| | | | |  
Db 553 CCAAGTGCACAACATGCTGTGTCTGTGTCCCGCCCAACAAGCAACCCAGCCACCGCA 612  
| | | | |  
QY 5265 TGAACGCTTGTGCTAATCTCCCAACGCGGCCCAAGCCTTGACAGCGCCCAAGCAAGT 5324  
| | | | |  
Db 613 TGAACGCTTGTGCTAATCTCCCAACGCGGCCCAAGCCTTGACAGCGCCCAAGCAAGT 672  
| | | | |  
QY 5335 CCCCCTCTCCCAAGAGAGTCCAAACAATTGACAAAACCAACCAACCAAGTCTGTCCG 5384  
| | | | |  
Db 673 CCCCCTCTCCCAAGAGAGTCCAAACAATTGACAAAACCAACCAACCAAGTCTGTCCG 732  
| | | | |  
QY 5385 AGCGGAGCGAGAACGCGGATCGAGAGCGGGAACGCGGATCGGAGCGGGAAGTCAATCC 5444  
| | | | |  
Db 733 AGCGGAGCGAGAACGCGGATCGAGAGCGGGAACGCGGATCGGAGCGGGAAGTCAATCC 792  
| | | | |  
QY 5445 TCAAGTCCACAACAAGCGGTGAGACGCAACCATCTGAGACCTGTGTACAGAGCAAGCA 5504  
| | | | |  
Db 793 TCAAGTCCACAACAAGCGGTGAGACGCAACCATCTGAGACCTGTGTACAGAGCAAGCA 852  
| | | | |  
QY 5505 GCGGAGAGAGCGGAGCAAGCGCGCGGGGTGGGGGAGAGCAAGCAAGCGCGCGCTCCCACT 5564  
| | | | |  
Db 853 GCGGAGAGAGCGGAGCAAGCGCGCGGGGTGGGGGAGAGCAAGCAAGCGCGCGCTCCCACT 912  
| | | | |  
QY 5565 CCAATGACCACAAGCACTGCGCCATCTCCCTCGAGACCAAGGATCCCTTCAGAGAGAG 5624  
| | | | |  
Db 913 CCAATGACCACAAGCACTGCGCCATCTCCCTCGAGACCAAGGATCCCTTCAGAGAGAG 972  
| | | | |  
QY 5625 CCAAGTGTCTTACAACAACAAGGATGAAGGATATCATCAACGCTGTGAGAGCCCAAGC 5684  
| | | | |  
Db 973 CCAAGTGTCTTACAACAACAAGGATGAAGGATATCATCAACGCTGTGAGAGCCCAAGC 1032  
| | | | |  
QY 5685 CCAAGTGTCTTACAACAACAAGGATGAAGGATATCATCAACGCTGTGAGAGCCCAAGC 5744  
| | | | |  
Db 1033 CCAAGTGTCTTACAACAACAAGGATGAAGGATATCATCAACGCTGTGAGAGCCCAAGC 1092  
| | | | |  
QY 5745 CTGCAACCACTGCGCCATGCGGCGGCAACCTGATGGGGTCTAACCCTAACCTCATGAGC 5804  
| | | | |  
Db 1093 CTGCAACCACTGCGCCATGCGGCGGCAACCTGATGGGGTCTAACCCTAACCTCATGAGC 1152  
| | | | |  
QY 5805 CCGTCTTGTCTGCGCAAGAGGCGCCCGGGGTGCGCGGCGAGAGCGGCCCGAGCAACA 5864  
| | | | |  
Db 1153 CCGTCTTGTCTGCGCAAGAGGCGCCCGGGGTGCGCGGCGAGAGCGGCCCGAGCAACA 1212  
| | | | |  
QY 5865 CCGGCAATGCTCTTCTGCGCAAGGCGCCCGGCTCGGGGTGAGAGCCGCTCTCTCC 5924  
| | | | |  
Db 1213 CCGGCAATGCTCTTCTGCGCAAGGCGCCCGGCTCGGGGTGAGAGCCGCTCTCTCC 1272  
| | | | |  
QY 5925 CCAAGCAAGGCTCGAGAGCGCGGCGCCCTAAGTGCCTCTGTCTCTGAGCAACGCAATG 5984  
| | | | |  
Db 1273 CCAAGCAAGGCTCGAGAGCGCGGCGCCCTAAGTGCCTCTGTCTCTGAGCAACGCAATG 1332  
| | | | |  
QY 5985 CCGGCAACCTGCGAAGAACTTGACCTGACACAGCGCAAGCCCGAGCCGCGGCGCAC 6044  
| | | | |  
Db 1333 CCGGCAACCTGCGAAGAACTTGACCTGACACAGCGCAAGCCCGAGCCGCGGCGCAC 1392  
| | | | |  
QY 6045 CTGCTCTGCGCTCGAGACCGGCAACCGGGAAGAACTCAAGTAAACCTTTTTCATCCAG 6104  
| | | | |  
Db 1393 CTGCTCTGCGCTCGAGACCGGCAACCGGGAAGAACTCAAGTAAACCTTTTTCATCCAG 1452  
| | | | |  
QY 6105 AACTGGAATCTCGGTTCTGAGGTATCAAGGAGAGCTACAGCCCGAAGGGGTGAGC 6164  
| | | | |  
Db 1453 AACTGGAATCTCGGTTCTGAGGTATCAAGGAGAGCTACAGCCCGAAGGGGTGAGC 1512  
| | | | |  
QY 6165 CCGTCAAGCCGTGTGAGCTCAACCAAGTCTGACCAAGCAAGAGGCTCCCAAGCACTGG 6224  
| | | | |

Db 1513 CCGTCAAGCCCTGTGAGCTACCCAGTGAACCAAGCAAGAGGCTCCCAAGCACTGG 1572  
| | | | |  
QY 6225 AAGAGCTGACAAGAGCCACCTGAGAGGGGAGCTGCGGCCCAAGAGCCAGGCGCCCTGA 6284  
| | | | |  
Db 1573 AAGAGCTGACAAGAGCCACCTGAGAGGGGAGCTGCGGCCCAAGAGCCAGGCGCCCTGA 1632  
| | | | |  
QY 6285 AGCTTGGCGGGAGAGCGCGCCACCTGCCACAACCTTGCGCGCTGCTGCTGAGAGCCAGCT 6344  
| | | | |  
Db 1633 AGCTTGGCGGGAGAGCGCGCCACCTGCCACAACCTTGCGCGCTGCTGCTGAGAGCCAGCT 1692  
| | | | |  
QY 6345 GGTCCAGCCCGCTGTCTCAAGACCGGCCCAAGGGGTAAAGGTCAACAGCGGGTGTCAACC 6404  
| | | | |  
Db 1693 GGTCCAGCCCGCTGTCTCAAGACCGGCCCAAGGGGTCAAGAGGTCAACAGCGGGTGTCAACC 1752  
| | | | |  
QY 6405 TGGCCCAAGCATCATGAGAGGTATCAACAAGACTACACCGGCAACCAAGCAAGCAGCAGC 6464  
| | | | |  
Db 1753 TGGCCCAAGCATCATGAGAGGTATCAACAAGACTACACCGGCAACCAAGCAAGCAGCAGC 1812  
| | | | |  
QY 6465 TCAAGCAACCCCTGCGCGCCCTCTAATCTCTTCCCTGGGGCCAGCTGCCGCTCTGG 6524  
| | | | |  
Db 1813 TCAAGCAACCCCTGCGCGCCCTCTAATCTCTTCCCTGGGGCCAGCTGCCGCTCTGG 1872  
| | | | |  
QY 6525 AACTTCGCGCCCAACCAAGTGAACCTTACCTTCGCGCCCGGCAACATGCTGCCGCGCC 6584  
| | | | |  
Db 1873 AACTTCGCGCCCAACCAAGTGAACCTTACCTTCGCGCCCGGCAACATGCTGCCGCGCC 1932  
| | | | |  
QY 6585 GTGGCTCCCGCCACAAGAGGGGCAAGAGGTCTCCAGAGCAAAAGAGTGGCT 6644  
| | | | |  
Db 1933 GTGGCTCCCGCCACAAGAGGGGCAAGAGGTCTCCAGAGCAAAAGAGTGGCT 1992  
| | | | |  
QY 6645 TGGGTGTGTGAGACGGAATTGAACCTGTGTCCCAACGAGGGCATGACGAGCCAG 6704  
| | | | |  
Db 1993 TGGGTGTGTGAGACGGAATTGAACCTGTGTCCCAACGAGGGCATGACGAGCCAG 2052  
| | | | |  
QY 6705 GGCACCTCCCGAGTGTGTGTACCGGCTGTGTACCGGATGGGGAAACAAGACGAGCCCA 6764  
| | | | |  
Db 2053 GGCACCTCCCGAGTGTGTGTACCGGCTGTGTACCGGATGGGGAAACAAGACGAGCCCA 2112  
| | | | |  
QY 6765 GGAAGATGGGCTCCAAGCTCCAGGCAACACAGGCAAGCGGCAAGCTTCTTCAAGCAGC 6824  
| | | | |  
Db 2113 GGAAGATGGGCTCCAAGCTCCAGGCAACACAGGCAAGCGGCAAGCTTCTTCAAGCAGC 2172  
| | | | |  
QY 6825 TGAACGAGAGCAACTCCGCAATGCTCAAGTCCAAGACCAAGAGATCAACAAGAGCTGA 6884  
| | | | |  
Db 2173 TGAACGAGAGCAACTCCGCAATGCTCAAGTCCAAGACCAAGAGATCAACAAGAGCTGA 2232  
| | | | |  
QY 6885 AACCCACAACCGGAATGAGCCTGAATCAATATCAGCCAGCCTGGGACGAGATCTTCA 6944  
| | | | |  
Db 2233 AACCCACAACCGGAATGAGCCTGAATCAATATCAGCCAGCCTGGGACGAGATCTTCA 2292  
| | | | |  
QY 6945 ATATGCCCCCATCAACCGGAACAGGCCCTTATGACCTATGAAGGCAAGCGGTCGAGAAC 7004  
| | | | |  
Db 2293 ATATGCCCCCATCAACCGGAACAGGCCCTTATGACCTATGAAGGCAAGCGGTCGAGAAC 2352  
| | | | |  
QY 7005 ATGCAAGCAACAATGAGGGCTGAGAGCCATTAATTGAAGGCACTCATGGGTAAATATG 7064  
| | | | |  
Db 2353 ATGCAAGCAACAATGAGGGCTGAGAGCCATTAATTGAAGGCACTCATG----- 2402  
| | | | |  
QY 7065 ACCAGTGGAAAGTCCCGCGCTCAGCGCAATGCTTTTAAACCTCTGAATGCCAGTG 7124  
| | | | |  
Db 2403 ----- 2402  
| | | | |  
QY 7125 CCAAGCTGCGCGCTGTATGCGCATTAACGCGTGTGAACGAGCGAGTGAACCAACACTCA 7184  
| | | | |  
Db 2403 ----- 2402  
| | | | |  
QY 7185 CCTGCGAGTGTGCGCGGAGGCAAGGTCCTGTGAGACCCAGCAAGCGGAAAAAGCA 7244  
| | | | |  
Db 2403 -----GTTGGCGCGGGAAGGCAAGGTCTGTGGAGAACCAAGCGGAAAAAGCA 2454  
| | | | |  
QY 7245 AGTCCCGGCGCGGCGCTGAGCTGTGGGAACCGGCAACCTCTGTCTTCAAGTCACT 7304  
| | | | |

Db	2455	AGTCCCCGGCCCGGGGCTGTGACTGTGGGGACCGGCAACCTCTGTCTCTCAATGCACT	2514
Qy	7305	CGAGGGAGACTGCACATCGCCGGACGCGCTTCACCAACCGCTGTGGAGGACAGCGCTT	7364
Db	2515	CGAGGGAGACTGCACACCGCGGACGCGCTCACCAACCGCTGTGGAGGACAGCGCTT	2574
Qy	7365	CGTCCGCAAGTTTCAACGCATTCCTCTCAACACCCCTATCATCGGCTGACAGCGGGTG	7424
Db	2575	CGTCCGCAAGTTTCAACGCATTCCTCTCAACACCCCTATCATCGGCTGACAGCGGGTG	2634
Qy	7425	TCATGGCTTCCCCACCCCAACCGGGCTCCCCCGGGGACGCGGCGCTTGGCTGGCCCC	7484
Db	2635	TCATGGCTTCCCCACCCCAACCGGGCTCCCCCGGGGACGCGGCGCTTGGCTGGCCCC	2694
Qy	7485	ACCAAGCCTGGGAGAGAGGAGCCCAAGCCATGCTCTGTGTCAGATGACAGACATCTCCG	7544
Db	2695	ACCAAGCCTGGGAGAGAGGAGCCCAAGCCATGCTCTGTGTCAGATGACAGACATCTCCG	2754
Qy	7545	ACACGAGTGACTCAGAACAGGGCGGGGGGGGGCGGCGGATBUEBSPERLUGTCAGGTCC	7604
Db	2755	ACACGAGTGACTCAGAACAGGGCGGGGGGGGGCGGCGGATBUEBSPERLUGTCAGGTCC	2806
Qy	7605	CAGGAGGCA CAGGAACGGCCCTCTCAGAGAGGGGGCGGCTGCGCATCCCCCAACCAAG	7664
Db	2807	CAGGAGGCA CAGGAACGGCCCTCTCAGAGAGGGGGCGGCTGCGCATCCCCCAACCAAG	2866
Qy	7665	AAGGAGCCCCCTGAAGTCGGCTTGGCGCTTCATCATCTGTCCGTCCAGAGCCGCGATCCTT	7724
Db	2867	AAGGAGCCCCCTGAAGTCGGCGCTTCATCATCTGTCCGTCCAGAGCCGCGATCCTT	2926
Qy	7725	GCCT 7728	
Db	2927	GCCT 2930	

RESULT 15  
US-10-174-014-13  
Sequence 13, Application US/10174014  
Publication No. US20040005292A1  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
APPLICANT: Susan M. Preter  
APPLICANT: Kenneth W. Doble  
TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION  
FILE REFERENCE: PFS-0012  
CURRENT APPLICATION NUMBER: US/10174,014  
CURRENT FILING DATE: 2002-06-17  
NUMBER OF SEQ ID NOS: 73  
SEQ ID NO 13  
LENGTH: 2930  
TYPE: DNA  
ORGANISM: H. sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (456) ... (2765)  
US-10-174-014-13

Query Match	Score	DB	Length
Best Local Similarly	30.6%	2618.6	2930
Matches 2786; Conservative	94.6%	Pred. No. 0	
	3	Mismatches	9
		Indels	146
		Gaps	2

QY	4785	CCCCGACAGCACCCTGCGCCGAGCACCACACCCCATCTCGCCCTATGACACTGC	4844
Db	133	CCCCGACAGCACCCTGCGCCGAGCACCACACCCCATCTCGCCCTATGACACTGC	192
QY	4845	TTCCGGGCGTGAATGCGCGTGAAGCTGTATCGCAGCCACATCCCTCGGCTTCGACCCCA	4904
Db	193	TTCCGGGCGTGAATGCGCGTGAAGCTGTATCGCAGCCACATCCCTCGGCTTCGACCCCA	252
QY	4905	CCTCCATATCCCCCGGCGCATTCCTCTGAGAGCAGACCGCGTGCCTATCTACTGCCCCGACAC	4966
Db	253	CCTCCATATCCCCCGGCGCATTCCTCTGAGAGCAGACCGCGTGCCTATCTACTGCCCCGACAC	312

QY	4965	TGGGCCCCAACCCCACTTACCCCGACCTGTATCCCACTTACCTTATCGCGGCTTACCCG	5024
Db	313	TGGCCCCCAACCCCACTTACCCCGACCTGTATCCCACTTACCTTATCGCGGCTTACCCG	372
QY	5025	AACGGCGGGCCTGTGAGAACCGGAGACCATCATCTAACTGACTATCATCTCGAGAGA	5084
Db	373	AACGGCGGGCCTGTGAGAACCGGAGACCATCATCTAACTGACTATCATCTCGAGAGA	432
QY	5085	TGCACCAACAACGGCCACCGCCATGGCCGACAGCTGATATGCTGAGGGGCTTCTGC	5144
Db	433	TGCACCAACAACGGCCACCGCCATGGCCGACAGCTGATATGCTGAGGGGCTTCTGC	492
QY	5145	CCCGGAGTCTCGCTGGCACTCAACTACGTGGGGGCCCCGAGGGCATATGACCTGT	5204
Db	493	CCCGGAGTCTCGCTGGCACTCAACTACGTGGGGGCCCCGAGGGCATATGACCTGT	552
QY	5205	CCCAATGCCACACCTGTCTGTCTGTCCCGACACACAGGACCCCAAGCCACCGCA	5264
Db	553	CCCAATGCCACACCTGTCTGTCTGTCCCGACACACAGGACCCCAAGCCACCGCA	612
QY	5265	TGGACCCGCTTGGCTTACTCTCCCAACCGGCCCAAGCCCTTCAAGAGCCGACAGAGCT	5324
Db	613	TGGACCCGCTTGGCTTACTCTCCCAACCGGCCCAAGCCCTTCAAGAGCCGACAGAGCT	672
QY	5325	CCCCACTCTCCCAAGAGGTTCCAAACAATTGACAAACCAACCAACGCTCTGTCCG	5384
Db	673	CCCCACTCTCCCAAGAGGTTCCAAACAATTGACAAACCAACCAACGCTCTGTCCG	732
QY	5385	AGCGGAGCGAGACCGGGATTCGAGACCGGACCGGGATTCGGAGCCGGAAAAATTCATCC	5444
Db	733	AGCGGAGCGAGACCGGGATTCGAGAGCGGGACCGGGATTCGGAGAAATTCATCC	792
QY	5445	TCAAGTCCACACGACGGTGGAGACGACCCATCTGGAAGCTGGATACAGAGAGAGA	5504
Db	793	TCAAGTCCACACGACGGTGGAGACGACCCATCTGGAAGCTGGATACAGAGAGAGA	852
QY	5505	GGCGCAGCAGCGGAGCAGCGGCGGGGATGGGGCACAAGCAGCGCGGCTCCCACT	5564
Db	853	GGCGCAGCAGCGGAGCAGCGGCGGGGATGGGGCACAAGCAGCGCGGCTCCCACT	912
QY	5565	CCCATGGCCACACAGCACTCGGCCATCTCCCTCGGAGACCCAGATGCGCTTCAAGAGAGAC	5624
Db	913	CCCATGGCCACACAGCACTCGGCCATCTCCCTCGGAGACCCAGATGCGCTTCAAGAGAGAC	972
QY	5625	CCAGTGTGCTTCAACAACAAGGCATGGAAGGATCATCAACGCTGTGAGGCCAGCAGC	5684
Db	973	CCAGTGTGCTTCAACAACAAGGCATGGAAGGATCATCAACGCTGTGAGGCCAGCAGC	1032
QY	5685	CCACGGTCTTGAAGTCCACTTCACTTCAACCGGTTGCGCCAGCTGCCACATTTCCAC	5744
Db	1033	CCACGGTCTTGAAGTCCACTTCACTTCAACCGGTTGCGCCAGCTGCCACATTTCCAC	1092
QY	5745	CTTGCACCAACTGCGCCACTGGGCGGCAACCTCGATGGGGTCTAACCTCATGAGAGC	5804
Db	1093	CTTGCACCAACTGCGCCACTGGGCGGCAACCTCGATGGGGTCTAACCTCATGAGAGC	1152
QY	5805	CCGTCTTGTGCTCCCAAGAGGCCCCCGGGTGGCCCGGACAGACCGGCCCGGACAGACA	5864
Db	1153	CCGTCTTGTGCTCCCAAGAGGCCCCCGGGTGGCCCGGACAGACCGGCCCGGACAGACA	1212
QY	5865	CCGGCCATGCTTCTGTGCCCAAGCCCCCAAGCCGCTTCGGGCTGGAGCCCGGCTTCTCC	5924
Db	1213	CCGGCCATGCTTCTGTGCCCAAGCCCCCAAGCCGCTTCGGGCTGGAGCCCGGCTTCTCC	1272
QY	5925	CCACCAAGGGCTGGAGACCCCGGCCCTTATGCTCTGTCTCTGAGCAGCCACATCG	5984
Db	1273	CCACCAAGGGCTGGAGACCCCGGCCCTTATGCTCTGTCTCTGAGCAGCCACATCG	1332
QY	5985	CCGCGACCCCTGCGAAGAACCTTGCACTTCAACGCGCAACCGGAGCCCGCGGCGCAC	6044
Db	1333	CCGCGACCCCTGCGAAGAACCTTGCACTTCAACGCGCAACCGGAGCCCGCGGCGCAC	1392

QY 6045 CTGCTCGGCTCGGACCGGAAAGACTCAAAAGTAAACCTTTTCATCCAGG 6104  
Db 1393 CTGCTCGGCTCGGACCGGAAAGACTCAAAAGTAAACCTTTTCATCCAGG 1452  
QY 6105 AACTGGAATCTCGTTCTCTGGGTTTACCAAGGAGAGCTACAGCCCGAAGGGGGTGGAGC 6164  
Db 1453 AACTGGAATCTCGTTCTCTGGGTTTACCAAGGAGAGCTACAGCCCGAAGGGGGTGGAGC 1512  
QY 6165 CCGTCAAGCCCTGTAGCTCAACCAAGTCTGACCAAGCAAGGGGCTCCCAAGCACTGG 6224  
Db 1513 CCGTCAAGCCCTGTAGCTCAACCAAGTCTGACCAAGCAAGGGGCTCCCAAGCACTGG 1572  
QY 6225 AAGAGCTGCAAGAGCCACTTGGAGGGGAGCTGGGCCCAAGCAAGCCCGCTTGA 6284  
Db 1573 AAGAGCTGCAAGAGCCACTTGGAGGGGAGCTGGGCCCAAGCAAGCCCGCTTGA 1632  
QY 6285 AGCTTGGGGGGAGGCGCGCCCACTCCCAACTGGGGCGGTGCTGAGAGCCAGCCCT 6344  
Db 1633 AGCTTGGGGGGAGGCGCGCCCACTCCCAACTGGGGCGGTGCTGAGAGCCAGCCCT 1692  
QY 6345 CGTCCAGCCCGCTGCTCCAGACCGGCCCAAGGGTCAAAAGTCAACAGGGGTGTCAACC 6404  
Db 1693 CGTCCAGCCCGCTGCTCCAGACCGGCCCAAGGGTCAAAAGTCAACAGGGGTGTCAACC 1752  
QY 6405 TGGCCCAAGCAATCATGTAGTCTATCAACAGAGCTACACCCGGCAACCAACAGCAGC 6464  
Db 1753 TGGCCCAAGCAATCATGTAGTCTATCAACAGAGCTACACCCGGCAACCAACAGCAGC 1812  
QY 6465 TCAGGCAACCCCTGGCGCGCCCTCTACTCTTCCCTGGGGCGAGCTGGCCCGCTGG 6524  
Db 1813 TCAGGCAACCCCTGGCGCGCCCTCTACTCTTCCCTGGGGCGAGCTGGCCCGCTGG 1872  
QY 6525 AACTCCGCGCGCCCAACCAAGTGACTCTTAACCTCCCGCCCGGACCAATGGTGGCCCGGCC 6584  
Db 1873 AACTCCGCGCGCCCAACCAAGTGACTCTTAACCTCCCGCCCGGACCAATGGTGGCCCGGCC 1932  
QY 6585 GTGGCTCCCCCAAGCAAGGAGGGGCAAGAGTCTTCAGAGCCAAACAGAGCTGGTCT 6644  
Db 1933 GTGGCTCCCCCAAGCAAGGAGGGGCAAGAGTCTTCAGAGCCAAACAGAGCTGGTCT 1992  
QY 6645 TGGGTTGTTGTTGAGGACGGTATTGAACCTGTGTCCCAACCGGAGGCAATGACGAGCCAG 6704  
Db 1993 TGGGTTGTTGTTGAGGACGGTATTGAACCTGTGTCCCAACCGGAGGCAATGACGAGCCAG 2052  
QY 6705 GGCATCTCCGAGAGTCTGTATCCGCTGCTGTACCGGAGTGGGGAACAGAGGACCA 6764  
Db 2053 GGCATCTCCGAGAGTCTGTATCCGCTGCTGTACCGGAGTGGGGAACAGAGGACCA 2112  
QY 6765 GCAGGATGGGCTCAAGTCTCCAGGCAACCAAGCCAGCCGCTTCTTCAAGCAAGC 6824  
Db 2113 GCAGGATGGGCTCAAGTCTCCAGGCAACCAAGCCAGCCGCTTCTTCAAGCAAGC 2172  
QY 6825 TGAACGAGAGCACTCCGCAATGTCAGTCCAGAGCAAGAGATCAACAGAGCTGA 6884  
Db 2173 TGAACGAGAGCACTCCGCAATGTCAGTCCAGAGCAAGAGATCAACAGAGCTGA 2232  
QY 6885 AACACCAACAACCGAATGAGCTGAATACAATATCAGCAGCCTGGGACGAGATCTTCA 6944  
Db 2233 AACACCAACAACCGAATGAGCTGAATACAATATCAGCAGCCTGGGACGAGATCTTCA 2292  
QY 6945 ATATGCCCCCAATCAACCGAACAAGGCTTATGACTATGAAAGCCAGGCGGTGACGAAAC 7004  
Db 2293 ATATGCCCCCAATCAACCGAACAAGGCTTATGACTATGAAAGCCAGGCGGTGACGAAAC 2352  
QY 7005 ATGCGAGCAACCAATGGGGCTGGAGGCAATTAAGAAAGGCACTCATGGTAAATATG 7064  
Db 2353 ATGCGAGCAACCAATGGGGCTGGAGGCAATTAAGAAAGGCACTCATG----- 2402  
QY 7065 ACCAGTGGAAAGTCCCGCGCTCAGGCGCAATGCTTTAAACCTCTGAATGCCAGTG 7124  
Db 2403 ----- 2402  
QY 7125 CCAGCCTGCGCGCTGTAATGCCATTAACCGCTGTGACGAGCAGAGTGAACACACTCA 7184

Db 2403 ----- 2402  
QY 7185 CCTGCGAGTGGCGGCGGAAAGGCCAAGTCTCTGGGAGAACCCAGCAGCCGAAAGGCA 7244  
Db 2403 -----GTTGGCGGCGGAAAGGCCAAGTCTCTGGGAGAACCCAGCAGCCGAAAGGCA 2454  
QY 7245 AGTCCCGGCCCCCGGAGCTTGGATCTGGGGAACCGGCAACCTCTGTCTTCAAGTGA 7304  
Db 2455 AGTCCCGGCCCCCGGAGCTTGGATCTGGGGAACCGGCAACCTCTGTCTTCAAGTGA 2514  
QY 7305 CGGAGGGAATCTGCAACCGCGGGAACCGCGCTCAACCAACCGGTGGAGGAGCAAGCCCT 7364  
Db 2515 CGGAGGGAATCTGCAACCGCGGGAACCGCGCTCAACCAACCGGTGGAGGAGCAAGCCCT 2574  
QY 7365 CGTCCGAGGTTCCAGGCAATTCCTCAACCCCTGATCATGGGCTGACAGCGGGTG 7424  
Db 2575 CGTCCGAGGTTCCAGGCAATTCCTCAACCCCTGATCATGGGCTGACAGCGGGTG 2634  
QY 7425 TCATGGCTTCCCAACCCCAACCGGCTCTCCCGGGGAGCGGGCCCTCGCTGGCCCC 7484  
Db 2635 TCATGGCTTCCCAACCCCAACCGGCTCTCCCGGGGAGCGGGCCCTCGCTGGCCCC 2694  
QY 7485 ACCAAGCTTGGGAGGAGAGGCCAAGCCACTGTCTGTGCAAGTACAGAGCACTTCCG 7544  
Db 2695 ACCAAGCTTGGGAGGAGAGGCCAAGCCACTGTCTGTGCAAGTACAGAGCACTTCCG 2754  
QY 7545 ACAGGAGTGACTCAGAACAGGCGGGGGGGGCGGCGGTGEBERSPERLUTGAGTCC 7604  
Db 2755 ACAGGAGTGACTCAGAACAGGCGGGGGGGGCGGCGGTGEBERSPERLUTGAGTCC 2806  
QY 7605 CAGCGAGCCACAGGAACGGCCCTGTGAGAGCGGGGCGGCTGCCAATCCCCAACGAAG 7664  
Db 2807 CAGCGAGCCACAGGAACGGCCCTGTGAGAGCGGGGCGGCTGCCAATCCCCAACGAAG 2866  
QY 7665 AAGGAGCCCTGAGTCCGCTGCGCTCAATCAATCTGTCCGTCCAGAGCCGGGATCTT 7724  
Db 2867 AAGGAGCCCTGAGTCCGCTGCGCTCAATCAATCTGTCCGTCCAGAGCCGGGATCTT 2926  
QY 7725 GCGT 7728  
Db 2927 GCGT 2930

Search completed: March 12, 2006, 12:21:16  
Job time : 5998 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2006, 09:17:38 ; Search time 1329 Seconds  
(without alignments)  
11463.868 Million cell updates/sec

Title: US-09-522-753-4  
Perfect score: 8564.8  
Sequence: 1 catgctcggtccacacagc.....caaaaaaaaaaaaaa 8571

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents NA:\*  
1: /cgn2\_6/prodata/1/ina/1 COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5 COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/H COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq:\*  
7: /cgn2\_6/prodata/1/ina/PP COMB.seq:\*  
8: /cgn2\_6/prodata/1/ina/RE COMB.seq:\*  
9: /cgn2\_6/prodata/1/ina/backfilseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8226.8	96.1	9053	3	US-09-976-594-306 Sequence 306, App
2	8230	7.4	7912	3	US-09-949-016-3804 Sequence 3804, App
3	630	7.4	7940	3	US-09-632-0338-1 Sequence 1, App11
4	629.6	7.4	7780	3	US-09-632-0338-2 Sequence 2, App11
5	311.2	3.6	312	3	US-09-513-999C-33415 Sequence 33415, A
6	142	1.7	1922	2	US-08-372-652-9 Sequence 9, App11
7	142	1.7	1922	6	PCT-US95-16311-9 Sequence 9, App11
8	107.8	1.3	3489	2	US-08-728-323A-1 Sequence 1, App11
9	107.8	1.3	3489	3	US-09-298-568-1 Sequence 1, App11
10	107.8	1.3	3489	3	US-09-410-399-1 Sequence 1, App11
11	107.8	1.3	3489	3	US-09-894-273-1 Sequence 1, App11
12	107.8	1.3	32207	2	US-08-770-379-20 Sequence 20, App1
13	107.8	1.3	32207	3	US-08-757-669A-20 Sequence 20, App1
14	107.8	1.3	32207	3	US-09-230-371A-20 Sequence 20, App1
15	104.6	1.2	187595	3	US-09-949-016-15546 Sequence 15546, A
16	96.2	1.1	1926	3	US-09-249-585A-2 Sequence 2, App11
17	96.2	1.1	1926	3	US-09-410-399-3 Sequence 3, App11
18	96.2	1.1	2580	3	US-09-050-863-2 Sequence 2, App11
19	96.2	1.1	2580	3	US-09-359-081-2 Sequence 2, App11
20	96.2	1.1	5452	2	US-09-130-114-1 Sequence 1, App11
21	96.2	1.1	8705	3	US-09-647-344A-14 Sequence 14, App1
22	96.2	1.1	9600	3	US-08-910-647-1 Sequence 1, App1
23	96.2	1.1	9600	3	US-09-620-925-1 Sequence 1, App1
24	96.2	1.1	10596	2	US-07-884-811-15 Sequence 15, App1

25	96.2	1.1	10596	2	US-07-885-971-15 Sequence 15, App1
26	96.2	1.1	10596	2	US-08-087-783A-15 Sequence 15, App1
27	96.2	1.1	10596	2	US-08-194-088B-15 Sequence 15, App1
28	96.2	1.1	10596	2	US-08-194-087-15 Sequence 15, App1
29	96.2	1.1	10596	6	PCT-US93-04648-15 Sequence 15, App1
30	96.2	1.1	16080	3	US-09-724-566A-48 Sequence 48, App1
31	96.2	1.1	16080	3	US-09-471-669A-48 Sequence 48, App1
32	94.8	1.1	127	3	US-09-680-420A-17 Sequence 17, App1
33	94.8	1.1	601	3	US-09-949-016-135614 Sequence 135614, A
34	93	1.1	94	3	US-09-513-999C-29269 Sequence 29269, A
35	92	1.1	1852	3	US-09-969-852-4 Sequence 4, App11
36	90	1.1	1926	3	US-09-249-585A-4 Sequence 4, App11
37	90	1.1	1931	2	US-09-130-114-2 Sequence 2, App11
38	89.8	1.0	7218	2	US-08-232-463-14 Sequence 14, App1
39	86.2	1.0	1995	2	US-08-425-069-3 Sequence 3, App11
40	86.2	1.0	1995	2	US-08-317-844B-3 Sequence 3, App11
41	85.4	1.0	16442	3	US-08-781-891-208 Sequence 208, App
42	85.4	1.0	16442	3	US-09-618-166-208 Sequence 208, App
43	82.8	1.0	767677	3	US-09-949-016-12147 Sequence 12147, A
44	82.8	1.0	767677	3	US-09-949-016-17361 Sequence 17361, A
45	82.4	1.0	2093	3	US-10-104-047-1666 Sequence 1666, Ap

#### ALIGNMENTS

RESULT 1  
US-09-976-594-306  
Sequence 306, Application US/09976594  
Patent No. 6673549  
GENERAL INFORMATION:  
APPLICANT: Furness, Michael  
APPLICANT: Buchinder, Jenny  
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
FILE REFERENCE: PA-0041 US  
CURRENT APPLICATION NUMBER: US/09/976, 594  
CURRENT FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: 60/240, 409  
PRIOR FILING DATE: 2000-10-12  
NUMBER OF SEQ ID NOS: 1143  
SOFTWARE: PERL Program  
SEQ ID NO 306  
LENGTH: 9053  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No. 6673549 898877.6  
NAME/KEY: unsure  
LOCATION: 2006, 2012  
OTHER INFORMATION: a, t, c, g, or other  
US-09-976-594-306

Query Match 96.1%; Score 8226.8; DB 3; Length 9053;  
Best Local Similarity 98.3%; Pred. No. 0;  
Matches 8451; Conservative 2; Mismatches 42; Indels 103; Gaps 10;

1 CATGCGGCTCCACACAGCTTGTGACAGACGTGAGGAGCCACTGAGCCCGCTAACCC 60  
519 CATGTGGGATCCACACACAGCTGTGGACAGACGTGAGGAGCCACTGAGCCCGCTAACCC 578  
61 GCGCCACAGCTTCTTACCCAGTGCAGATCGCCGAGGACACAGGAGCGTGGGCTCCT 120  
579 GCGCCACAGCTTCTTACCCAGTGCAGATCGCCGAGGACACAGGAGCGTGGGCTCCT 638  
121 GAGATACAGACCACTCCCGGACATATGCTCCCACTGCGCGGGCTCATATCCA 180  
639 GAGATACAGACCACTCCCGGACATATGCTCCCACTGCGCGGGCTCATATCCA 698  
181 GCGCCAGCGGAGGAGCGCTCTGCTGTCTGATTTCAGCCCGGAGATGACGCTCCCA 240  
699 GCGCCAGCGGAGGAGCGCTCTGCTGTCTGATTTCAGCCCGGAGATGACGCTCCCA 758



QY 241 GAGACTCCACTGCGGCAGAGTCCCACTCATACCTGCCGAGCTGGGAACTCAGAGT 300  
Db 759 GAGACTCCACTGCGGCAGAGTCCCACTCATACCTGCCGAGCTGGGAACTCAGAGT 818  
QY 301 GAGACTTCATGAAACCAAGCCCTCCGGTGAAGCTGCTGCTGACCCCTGCTCTCGAC 360  
Db 819 GAGACTTCATGAAACCAAGCCCTCCGGTGAAGCTGCTGCTGACCCCTGCTCTCGAC 878  
QY 361 GTCAACCCCTGCTGCGCAAGCGGCAGAGCTGCGGATCTGAAAGACTCAACCAAGACCGTAG 420  
Db 879 GTCAACCCCTGCTGCGCAAGCGGCAGAGCTGCGGATCTGAAAGACTCAACCAAGACCGTAG 938  
QY 421 CTTGACGGGCAAGCTGGAACCGGTGTCTTCCCGCCAGCCCCCGCAACTGAACCTTGAGCT 480  
Db 939 CTTGACGGGCAAGCTGGAACCGGTGTCTTCCCGCCAGCCCCCGCAACTGAACCTTGAGCT 998  
QY 481 GAGAGTGTGCGCGCAAGCGGTGTCTTCCCGCCAGCCCCCGCAACTGAACCGGTGGA 540  
Db 999 GAGAGTGTGCGCGCAAGCGGTGTCTTCCCGCCAGCCCCCGCAACTGAACCGGTGGA 1058  
QY 541 CCGAGAGATCAACATGATGAGCAGACAGATCTTAAGCTGAAGAAAGACAGCAACAGT 600  
Db 1059 CCGAGAGATCAACATGATGAGCAGACAGATCTTAAGCTGAAGAAAGACAGCAACAGT 1118  
QY 601 GAGAGAGAGAGCTGCGCAAGCGGCAGAGCTGGAAGACCCGTGTCAACGCGCGCCATGGA 660  
Db 1119 GAGAGAGAGAGCTGCGCAAGCGGCAGAGCTGGAAGACCCGTGTCAACGCGCGCCATGGA 1178  
QY 661 GTGGAAGACCGGCAGAGCTGATGACATCATCTACACAGAAACCGGAAGAAAGCTGAAGC 720  
Db 1179 GTGGAAGACCGGCAGAGCTGATGACATCATCTACACAGAAACCGGAAGAAAGCTGAAGC 1238  
QY 721 TGCAATCATGATTTGGAAGGCTTGGGCGCCAGGTGAGCTGCCCTGTACAACAGCC 780  
Db 1239 TGCAATCATGATTTGGAAGGCTTGGGCGCCAGGTGAGCTGCCCTGTACAACAGCC 1298  
QY 781 CTCGCAACCCGCGCAGTATCATGAGAACATCAAAATTAACCAAGGATGCGGAAGAGCT 840  
Db 1299 CTCGCAACCCGCGCAGTATCATGAGAACATCAAAATTAACCAAGGATGCGGAAGAGCT 1358  
QY 841 AATCTTTGATCTTCAAGAGAGAGAAATCAAGCTTGGAAACATTTGGAAGCTTTGCGCA 900  
Db 1359 AATCTTTGATCTTCAAGAGAGAGAAATCAAGCTTGGAAACATTTGGAAGCTTTGCGCA 1418  
QY 901 GCGCTATGACCAAGCTCATGAGAGGCTTGGAAATAAAGGTGAGCCATCGAATAACACC 960  
Db 1419 GCGCTATGACCAAGCTCATGAGAGGCTTGGAGAAAGAGGTGAGCCATCGAATAACACC 1478  
QY 961 GCGCGCGCGGCGCAAGAGAGAGAGGTGCGAGTACTACGAAAACAGTTCCCTGAGAT 1020  
Db 1479 GCGCGCGCGGCGCAAGAGAGAGAGGTGCGAGTACTACGAAAACAGTTCCCTGAGAT 1538  
QY 1021 CCGGAAGACCGCGAGCTGACAGAGACGCAATGACAGAGGTGGGCGACGGGGGCAAGTGG 1080  
Db 1539 CCGGAAGACCGCGAGCTGACAGAGACGCAATGACAGAGGTGGGCGACGGGGGCAAGTGG 1598  
QY 1081 GCTGTCCATGTGCGCGCGCGGACGAGACGAGAGTGTCAAGATCATGTGATGAGCTCTTC 1140  
Db 1599 GCTGTCCATGTGCGCGCGCGGACGAGACGAGAGTGTCAAGATCATGTGATGAGCTCTTC 1658  
QY 1141 AGAGCAGAGAGAACTTGAAGAGCAGATGCGCAGCTGGCGGTGATCCCGCCATGCTGTA 1200  
Db 1659 AGAGCAGAGAGAACTTGAAGAGCAGATGCGCAGCTGGCGGTGATCCCGCCATGCTGTA 1718  
QY 1201 CGAGCTGACCAAGCGGCTCAATGAATTCATCAACATGAACGGGCTTTATGGCGGACCCAT 1260  
Db 1719 CGAGCTGACCAAGCGGCTCAATGAATTCATCAACATGAACGGGCTTTATGGCGGACCCAT 1778  
QY 1261 GAAGGTGATCAAAAGACCGCAGAGTATGAAATGTGAGTGAGCAGAGAGAGAGACCTT 1320  
Db 1779 GAAGGTGATCAAAAGACCGCAGAGTATGAAATGTGAGTGAGCAGAGAGAGAGACCTT 1838  
QY 1321 CCGGAGAGAGTTTCAATGACAGCATCCAGAGAACTTTGGCTGATGCAATTCCTGGAAG 1380

Db 1839 CCGGAGAGAGTTTCAATGACAGCATCCAGAACTTTGGCTGATGCAATTCCTGGAAG 1898  
QY 1381 GAAGCAGTGGCTGAGTGGCTCTTACTACTGCTGAAGAAATGAGAACTTATA 1440  
Db 1899 GAAGCAGTGGCTGAGTGGCTCTTACTACTGCTGAAGAAATGAGAACTTATA 1958  
QY 1441 GAGCTGTGAGACGAGACTATCGCGCGCGGCAAGAGCCAGCAGCAACA--CAGCA 1497  
Db 1959 GAGCTGTGAGACGAGACTATCGCGCGCGGCAAGAGCCAGCAGCAACAGCAGCA 2018  
QY 1498 GAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1557  
Db 2019 GAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2078  
QY 1558 GAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1617  
Db 2079 GAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2138  
QY 1618 GAACGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1677  
Db 2139 GAACGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2198  
QY 1678 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1737  
Db 2199 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2258  
QY 1738 AGGCGCATCAACCGCTCAATGAGCTAATGAGGCGCAAGCAGAGAGAGAGAGAGAGAGAG 1797  
Db 2259 AGGCGCATCAACCGCTCAATGAGCTAATGAGGCGCAAGCAGAGAGAGAGAGAGAGAGAG 2318  
QY 1798 GCAGAGCGCGAGAGCTGAGCTCATGAGCTGAATGAGAGTTCTGCTGGAACAGAAAGAG 1857  
Db 2319 GCAGAGCGCGAGAGCTGAGCTCATGAGCTGAATGAGAGTTCTGCTGGAACAGAAAGAG 2378  
QY 1858 AATGAAACAGCCAGAGAAAGTCTCTTGAACACGAGCGGCACTGTGGGCTGAGCGCCG 1917  
Db 2379 AATGAAACAGCCAGAGAAAGTCTCTTGAACACGAGCGGCACTGTGGGCTGAGCGCCG 2438  
QY 1918 GATGTTGGGCTCAAGACTGATGAGCAGTGTGAAGACTTACTCACTCACTCAAGAGAG 1977  
Db 2439 GATGTTGGGCTCAAGACTGATGAGCAGTGTGAAGACTTACTCACTCACTCAAGAGAGAG 2498  
QY 1978 GCAGAACTTCATGAGACTTTGACAGCAGCAGAGCTGGAATGAGAAAGAGAGAGAGAGAG 2037  
Db 2499 GCAGAACTTCATGAGACTTTGACAGCAGCAGAGCTGGAATGAGAAAGAGAGAGAGAGAG 2558  
QY 2038 GCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2096  
Db 2559 GCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2618  
QY 2097 TGGTGAAGATGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2156  
Db 2619 TGGTGAAGATGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2678  
QY 2157 AGGAGCTGAAGCCTTACATGCTCTTGGAAATGAGGTGCCAGAGGGGAATGCACTGCGC 2216  
Db 2679 AGGAGCTGA----- 2688  
QY 2217 CAGCCACTGTCAACACAGCTTCAGACACCGAGAGCATCCCTCTCTCACTAGAGCGG 2276  
Db 2689 -AGCCACTGTCAACACAGCTTCAGACACCGAGAGCATCCCTCTCTCACTAGAGCGG 2747  
QY 2277 CCAAGGACACAGGGAGAGATGGGGCCAGAGCCCAAGCCCACTGGGGCGCCAG--GGGCA 2335  
Db 2748 CCAAGGACACAGGGAGAGATGGGGCCAGAGCCCAAGCCCACTGGGGCGCCAG--GGGCA 2807  
QY 2336 CCCCCAGGCCACCAACCCCAACAGAGAGATTCCTCGGGGCCCACTTGAAGCCACCCG 2395  
Db 2808 CCCCCAGGCCACCAACCCCAACAGAGAGATTCCTCGGGGCCCACTTGAAGCCACCCG 2867  
QY 2396 GCTCTGAAGCACCAGAGCCCTTACGCCCCCAAGCAACCCCACTGCGCTCTGACCT 2455

D	b	2868	GCCTTGAAGCCACCGGAGCCCTTACGCCCCCAACGACACCCCATGCGCTTTCGACCT	2927
O	y	2456	CCTCTGTGTGTCCCAAGAGAGAGAGAGAGACCGCAGCAGCGCCCTCCAGTGAAG	2515
D	b	2928	CCTCCTGTGTGCCCAAGAGAGAGAGAGAGACCGCAGCAGCGCCCTCCAGTGAAG	2987
O	y	2516	GAAGGGAGAGACGAAAGCCCCCGCGCTGAGAGCTTGCAGTGAGACAGGGAGGCC	2575
D	b	2988	GAAGGGAGAGACGAAAGCCCCCGCGCTGAGAGCTTGCAGTGAGACAGGGAGAGCC	3047
O	y	2576	GAGGAGCCCTGTCAAGAGGAGTGTGACCGGAGGAAGCCGAGGGAGGGGGCCGCAAGGGCAAG	2635
D	b	3048	GAGGAGCCCTGTCAAGAGGAGTGTGACCGGAGGAAGCCGAGGGAGGGGGCCGCAAGGGCAAG	3107
O	y	2636	GACCGGAGAGCGCGCTGAGAGCCACCGCCGAGGGGGCGCTCAAGAGCAGAGAAAGAGAGGC	2695
D	b	3108	GACCGGAGAGCGCGCTGTGAGGCCACCGGCCGAGGGGGCGCTCAAGAGCAGAGAAAGAGAGGC	3167
O	y	2696	GAGGAGCCGACAGGCCCAACATGACCAAGAGCTTGCGGGCGCCCCCAAGAGCAGGACTCCAGT	2755
D	b	3168	GAGGAGCCGACAGGCCCAACATGACCAAGAGCTTGCGGGCGCCCCCAAGAGCAGGACTCCAGT	3227
O	y	2756	GCTACCTGTACATGACAGAGAGTGTGATGAGGCCCGAGGGGGCGCCAGAAACCGGCTCTG	2815
D	b	3228	GCTACCTGTGTGACAGAGAGTGTGATGAGGCCCGAGGGGGCGCCAGAAACCGGCTCTG	3287
O	y	2816	TCCCAAGAGGCCACGACTCTTACACCCCGCATGAGCCGCCGGGCAATATGCTTCAACCCAG	2875
D	b	3288	TCCCAAGAGGCCACGACTCTTCAACCCCGCATGAGCCGCCGGGCAATATGCTTCAACCCAG	3347
O	y	2876	AAGCCACTGTGACCTGAAAGCAGCTGAAAGCAGGAGCGAGCGCTGCAATCCCCCATTCAGGTC	2935
D	b	3348	AAGCCACTGTGACCTGAAAGCAGCTGAAAGCAGGAGCGAGCGCTGCAATCCCCCATTCAGGTC	3404
O	y	2936	ACCAAAAGTCTATGAGCCCCCCCCGGAGAGACGACACTCCACCAAGCCAGCTCCCCAG-C	2994
D	b	3405	ACCAAAAGTCTATGAGCCCCCCCCGGAGAGACGACACTCCACCAAGCCAGCTCCCCAGAC	3464
O	y	2995	CCCAACCGCACCGGAAACCTGTACAGCGGAGAGGAGCGCCCTCAGACAGCTTGACAGAG	3054
D	b	3465	CCCAACCGCACCGGAAACCTGTACAGCGGAGAGGAGCGCCCTCAGACAGCTTGACAGAG	3524
O	y	3055	CCCCCGGGGCAAGAGCAGAGACCCCGGACCCCGCGCAGCAGAGG-----	3100
D	b	3525	CCCCCGGGGCAAGAGCAGAGACCCCGGACCCCGCGCAGCAGAGGAGAGAGGCTGT	3584
O	y	3101	-----GCCTTTCGAGCCGAGGCCCAGAGGCTGTGAGGAGACCCCTTGTCTGAC	3150
D	b	3585	GTTCTTTCGAGCCCTTTCGAGGCCGAGAGGCTGTGAGGAGACCCCTTGTCTGAC	3644
O	y	3151	TTTCCGGGCGTGGCCCTTCCCGCGTGGCCCGCCCGGTGAAGGTATCAAGGCTTCCCGCATGCGCC	3210
D	b	3645	TTTCCGGGCGTGGCCCTTCCCGCGTGGCCCGCCCGGTGAAGGTATCAAGGCTTCCCGCATGCGCC	3704
O	y	3211	GAACCCCTCAGACCTTTCCTACGCTGTCACTGTGTACCCACTGCGCCCTGGGCTCCATGA	3270
D	b	3705	GAACCCCTCAGACCTTTCCTACGCTGTCACTGTGTACCCACTGCGCCCTGGGCTCCATGA	3764
O	y	3271	CACGTCCCGGCGGTCGTGCGGGGCCACCCACCATCTTCCAAACCGGCTTCCCTCATCTC	3330
D	b	3765	CACGTCCCGGCGGTCGTGCGGGGCCACCCACCATCTTCCAAACCGGCTTCCCTCATCTC	3824
O	y	3331	CTTGTGCCAAGACCCCAAGGCTTCTGTGAGAGCAATATGTTGCCATTTCCCAAGAAATGT	3390
D	b	3825	CTTGTGCCAAGACCCCAAGGCTTCTGTGAGAGCAATATGTTGCCATTTCCCAAGAAATGT	3884
O	y	3391	GGTTCAGGCTCAGAGTCGCTACTAGAGCATGTCCAAAGGCGCCGGTGTGGGCTCTGTACAT	3450
D	b	3885	GGTTCAGGCTCAGAGTCGCTACTAGAGCATGTCCAAAGGCGCCGGTGTGGGCTCTGTACAT	3944
O	y	3451	GGGGCTGTCCCTGTCCCATGAGCCCAAAAGCTGTGACCTTTCAGCGAGTGAAGCAG	3510
D	b	3945	GGGGCTGTCCCTGTCCCATGAGCCCAAAAGCTGTGACCTTTCAGCGAGTGAAGCAG	4004

OY	3511	GCAGCTGTCCCCAAGCGGGGCCACAGGCTGGGCGCACCCGGAAGAGCTTGGGGGGTGTCCCAACGCCCA	3570
Db	4005	GCAGCTGTCCCCAAGCGGGGCCACAGGCTGGGCGCACCCGGAAGAGCTTGGGGGGTGTCCCAACGCCCA	4066
OY	3571	GGAGGCGCTCCGTGCTGAGAAGGGAACAGCTCTGGGCTCACTTCGGGGCGGAAGCATCACCA	3638
Db	4065	GGAGGCGCTCCGTGCTGAGAAGGGAACAGCTCTGGGCTCACTTCGGGGCGGAAGCATCACCA	4122
OY	3631	AGGCATTCCACGACAACGGGTGGCCCTGGACAGGCGCATCAACATACCGGGCTCCATCAC	3697
Db	4125	AGGCATTCCACGACAACGGGTGGCCCTGGACAGGCGCATCAATCCCCGGCTCCATCAC	4184
OY	3691	CCAGGGCACCGCAGCTGACGTCCTGTACAAAGGGACCATCAACAGGATCATTCGGGAGGA	3755
Db	4185	CCAGGGCACCGCAGCTGACGTCCTGTACAAAGGGACCATCAACAGGATCATTCGGGAGGA	4244
OY	3751	CAGCCCGAGTCTGCTTGGACCGCGGCCGGGAAGAACGCTGGCCAAAGGGCAAGTCATCTA	3810
Db	4245	CAGCCCGAGTCTGCTTGGACCGCGGCCGGGAAGAACGCTGGCCAAAGGGCAAGTCATCTA	4304
OY	3811	CGAAGGCAGAAAGGGCCACGTCCTTGTCTTATGAGGGTGGCATGTCTGTGACCCAGTGGCT	3877
Db	4305	CGAAGGCAGAAAGGGCCACGTCCTTGTCTTATGAGGGTGGCATGTCTGTGACCCAGTGGCT	4366
OY	3871	CAAAGAGGACGGGACGAAGCAGCTGAAAGACCCCCCATGAGACGGCGGCCCAAGCGAC	3933
Db	4365	CAAAGAGGACGGGACGAAGCAGCTGAAAGACCCCCCATGAGACGGCGGCCCAAGCGAC	4422
OY	3931	CTATGACATGATGAGAGGGCGCGGTGGGAGAGCAATCTCTCAGCAGCAGATCGAAGTCT	3990
Db	4425	CTATGACATGATGAGAGGGCGCGGTGGGAGAGCAATCTCTCAGCAGCAGATCGAAGTCT	4484
OY	3991	CATGGGCGCTGGCCATCCCGCCGGAGCGACAACGCCCCCAACACTCAAAAGAGCAGACCA	4055
Db	4485	CATGGGCGCTGGCCATCCCGCCGGAGCGACAACGCCCCCAACACTCAAAAGAGCAGACCA	4544
OY	4051	CATCCGGGGGTCACTACACAAAGGGAATCCCTCGGTCTTACCTGAGAGGCAAGAGACTA	4110
Db	4545	CATCCGGGGGTCACTACACAAAGGGAATCCCTCGGTCTTACCTGAGAGGCAAGAGACTA	4604
OY	4111	CTTGGCTTGGGAGGGCCAAAGTCTCTTAAAGCGGGAAGGGACAAGCTTCCGCCCAACCGCTCT	4177
Db	4605	CTTGGCTTGGGAGGGCCAAAGTCTCTTAAAGCGGGAAGGGACAAGCTTCCGCCCAACCGCTCT	4666
OY	4171	ACGGGACCTTGACGAGGCGCTACAGAAGCAGAGGCGCTGGAGCCCTCGTAAGTGAAGCGGAC	4233
Db	4665	ACGGGACCTTGACGAGGCGCTACAGAAGCAGAGGCGCTGGAGCCCTCGTAAGTGAAGCGGAC	4722
OY	4231	CCATGAGGGGCTGGTGGCCACAGGTGAAGAGAGGCGGCGGCTTCATTCATGAGATCCGCG	4290
Db	4725	CCATGAGGGGCTGGTGGCCACAGGTGAAGAGAGGCGGCGGCTTCATTCATGAGATCCGCG	4784
OY	4291	CGAAGGACTGTGGGACACGCCCGAGCTGGCCCTGGGCCCCCGGGCGGCTCAAGAGAGGGCTC	4355
Db	4785	CGAAGGACTGTGGGACACGCCCGAGCTGGCCCTGGGCCCCCGGGCGGCTCAAGAGAGGGCTC	4844
OY	4351	CATGACCGCAGGGCACCCCGCTCAAGTACGACACCGGCGGCTGCACACTGTGCTCCAAAA	4410
Db	4845	CATGACCGCAGGGCACCCCGCTCAAGTACGACACCGGCGGCTGCACACTGTGCTCCAAAA	4904
OY	4411	GCACGACGTACGCTCCCTCATTCGGAGGCCCGGCGCGACGTTCCACCCGTGCACCCGCT	4470
Db	4905	GCACGACGTACGCTCCCTCATTCGGAGGCCCGGCGCGACGTTCCACCCGTGCACCCGCT	4966
OY	4471	GGATGTATGAGCGGACGCGCGGGGACCTGGAACGTGGCTGTACAGAGAGAGACTGAAAG	4538
Db	4965	GGATGTATGAGCGGACGCGCGGGGACCTGGAACGTGGCTGTACAGAGAGAGACTGAAAG	5024
OY	4531	CGGGCCAGAGGACCGCCAGACAGCTCGGGGGGCTCATTTGCGGCGGCGCCCGCTCATTTGT	4590
Db	5025	CGGGCCAGAGGACCGCCAGACAGCTCGGGGGGCTCATTTGCGGCGGCGCCCGCTCATTTGT	5084

QY	4591	CCCTGAGCTGGGTAAAGCCGGGCGAGAGCCCTGACCTAATGAGACAACGGGGGACCCCTT	4650
Db	5085	GCTTAGCTGGGGCAAGCCGGCGAGAGCCCTTAACCTATAGGACAACGGGGGACCCCTT	5144
QY	4651	TGCCGGACCTCTCCACGAGGGTTGGCCCTGATGCATGCGGGAGGCCACCGCGCGCTTGCA	4710
Db	5145	TGCCGGACCTCTCCACGAGGGTTGGCCCTGATGCACACGGGGAGGCCACCGCGCGCTTGCA	5204
QY	4711	GGAGGGGAGCCTTTTCTGTCAGCAAGGCATCCCAAGACCTGAAGGCTGACGTCCTCG	4770
Db	5205	GGAGGGGAGCCTTTTCTGTCAGCAAGGCATCCCAAGGCTGACGTCGAGCCTCG	5264
QY	4771	TGAGATGCGCAAGTCCCGGCAACAGACCGGTGCCCGAGACCAACCAACCCCATCTGGCC	4830
Db	5265	TGAGATGCGCAAGTCCCGGCAACAGACCGGTGCCCGAGACCAACCAACCCCATCTGGCC	5324
QY	4831	CTATGAGCACTGTCGTTGCGGGCGGTAGTGGGTGGACCTGTATGAGACCAATCCCCCT	4890
Db	5325	CTATGAGCACTGTCGTTGCGGGCGGTAGTGGGTGGACCTGTATGAGACCAATCCCCCT	5384
QY	4891	GGCTTTCGACCCCACTTCATATCCCGCGGATCCCTCTGGACGACCGGCTGCTACTA	4950
Db	5385	GGCTTTCGACCCCACTTCATATCCCGCGGATCCCTCTGGACGACCGGCTGCTACTA	5444
QY	4951	CTGSCCCCGAACCCTGCGGCCCAACCCCACTTAACCGGACCTGTATACCAACCTTAACCTAT	5010
Db	5445	CTGSCCCCGAACCCTGCGGCCCAACCCCACTTAACCGGACCTGTATACCAACCTTAACCTAT	5504
QY	5011	CCGGGCGCTAACCCGACACAGCGGGCGCTGGAGAACCGGACACCATTCATATGACTACAT	5070
Db	5505	CCGGGCGCTAACCCGACACAGCGGGCGCTGGAGAACCGGACACCATTCATATGACTACAT	5564
QY	5071	CACCTTCGACAGATGACCAACAACCGGCAACCGGCATGGCCCAAGCGAGCTGATATGCT	5130
Db	5565	CACCTTCGACAGATGACCAACAAGCGGCACACGCACATGGCCCAAGCGAGCTGATATATGCT	5624
QY	5131	GAGGGGCTCTTCGCCCCCGGAGTCTCTCGTGGCATCAACTACGTCGGGGTCCCCGAGG	5190
Db	5625	GAGGGGCTCTTCGCCCCCGGAGTCTCTCGTGGCATCAACTACGTCGGGGTCCCCGAGG	5684
QY	5191	CATCATCGAACCCTGACCCCAAGTGCACACTGCTGTCGTGCCCCCGACACCGAGGAC	5250
Db	5685	CATCATCGAACCCTGACCCCAAGTGCACACTGCTGTCGTGCCCCCGACACCGAGGAC	5744
QY	5251	CCAGCGCACCGCACATGACACGCGCTTGCTTACCTCCACCGCGGCCACGCCCTTACAGAG	5310
Db	5745	CCAGCGCACCGCACATGACACGCGCTTGCTTACCTCCACCGCGGCCACGCCCTTACAGAG	5804
QY	5311	CCGCGCACAGCACTCCCATCTTCCGCCAGGAGTTCACAACAATTGACAAAACCAACAC	5370
Db	5805	CCGCGCACAGCACTCCCATCTTCCGCCAGGAGTTCACAACAATTGACAAAACCAACAC	5864
QY	5371	CAGGTCCTCGTCCGAGCGGGAGCGAGACCCGGAGTTCAGAGGGGACCGGGATTCGGGAGCG	5430
Db	5865	CAGGTCCTCGTCCGAGCGGGAGCGAGACCCGGAGTTCAGAGGGGACCGGGATTCGGGAGCG	5924
QY	5431	GGAAAAGTCATCTCTCACTGTCACACACGACCGGTGGAGCACCAACCATCTTGAGACCTGCG	5490
Db	5925	GGAAAAGTCATCTCTCACTGTCACACACGACCGGTGGAGCACCAACCATCTTGAGACCTGCG	5984
QY	5491	TACAGAGCAGACAGCGGCGACAGCGGCGTACAGCGCGCGGGGTGGGGGCGACAGACGCG	5550
Db	5985	TACAGAGCAGACAGCGGCGACAGCGGCGTACAGCGCGCGGGGTGGGGGCGACAGACGCG	6044
QY	5551	CCCGGCTCCCACTCCATGSCCAACAGCACTCCGCCATCTCCCTCGGACCCAGAGATGC	5610
Db	6045	CCCGGCTCCCACTCCATGSCCAACAGCACTCCGCCATCTCCCTCGGACCCAGAGATGC	6104
QY	5611	CCTTCAGCAGAGACCAAGTGTGCTTCAACAACAGGACATGAAGGATATCATCACCGCTGT	5670
Db	6105	CCTTCAGCAGAGACCAAGTGTGCTTCAACAACAGGACATGAAGGATATCATCACCGCTGT	6164
QY	5671	GGAGCCCGACAGGCCACAGTCTTG-----AGTTCACCTTCACCTTCCTACCCGT	5721

Db	6165	GGAGCCAGACGCGCCACCGTCTCTGAGGTGGGCAAGTCCACTCCACTCTTCAGCCGT	6224
Qy	5722	TCGCGCAGCTGCCACATTCCACTTGCACCACTGCCCCACTGGGGGCGACCTTGATGG	5781
Db	6225	TCGCGCGGCTGCCACATTCCACTTGCACCACTGCCCCACTGGGGGCGACCTTGATGG	6284
Qy	5782	GGTCTACCTTACCTCTTATGGAAGCCCGCTTTGGCTGGCCAGAGAGGCCCGCGGTGGCCG	5841
Db	6285	GGTCTACCTTACCTCTTATGGAAGCCCGCTTTGGCTGGCCAGAGAGGCCCGCGGTGGCCG	6344
Qy	5842	GCCAGAGCGGCGCCCGAGACAGACACCGGCGATGCTTCTCTGCGCAAGGCCCGACCGCGTC	5901
Db	6345	GCCAGAGCGGCGCCCGAGAGACACCGGCGATGCTTCTCTGCGCAAGGCCCGACCGCGTC	6404
Qy	5902	CGGCTGGAGCCCGCGCTCTCTCCCGACAGAGGGCTTGGAGCCCGCGCCCTTAGTGGCTTC	5961
Db	6405	CGGCTGGAGCCCGCGCTCTCTCCCGACAGAGGGCTTGGAGCCCGCGCCCTTAGTGGCTTC	6464
Qy	5962	TGTCTCTGGCGACAGCCACCATGCGCCCGACCCCTGCGAAGAACCTTGCACCTCACACGCG	6021
Db	6465	TGTCTCTGGCGACAGCCACCATGCGCCCGACCCCTGCGAAGAACCTTGCACCTCACACGCG	6524
Qy	6022	CAGCCCGGACCCCGCGCGCGCACCTGCTCGGCTCTGAGCCGCGCACCGGGAAAAAGCTCA	6081
Db	6525	CAGCCCGGACCCCGCGCGCGCACCTGCTCGGCTCTGAGCCGCGCACCGGGAAAAAGCTCA	6584
Qy	6082	AAGTAAACCTTTTTCATCCAGGAACTGAACTCCGTTCTTGGGTTACACGCGACAG	6141
Db	6585	AAGTAAACCTTTTTCATCCAGGAACTGAACTCCGTTCTTGGGTTACACGCGACAG	6644
Qy	6142	CTACAGCCCGCGAAGGGGTGAGGCGCGCTGACGCTGTGAGCTACCCAGTCTACCCACGA	6201
Db	6645	CTACAGCCCGCGAAGGGGTGAGGCGCGCTGACGCTGTGAGCTACCCAGTCTACCCACGA	6704
Qy	6202	CAAGGGGCTCCCAAGGACCTGGAAAGCTCGACAAAGACCACTGGAAGGGGAGCTGGG	6261
Db	6705	CAAGGGGCTCCCAAGGACCTGGAAAGCTCGACAAAGACCACTGGAAGGGGAGCTGGG	6764
Qy	6262	GCCCAAGCAGCCAGGCGCCCGTGAAGCTTGGCGGGAGGCGCGCCACCTTCCACACTGGG	6321
Db	6765	GCCCAAGCAGCCAGGCGCCCGTGAAGCTTGGCGGGAGGCGCGCCACCTTCCACACTGGG	6824
Qy	6322	GCGGTGCTTGAGAGCCAGGCTCTGTTCCAGCCCGCTGTTCCAGACCGCCCAAGGGGTCA	6381
Db	6825	GCGGTGCTTGAGAGCCAGGCTCTGTTCCAGCCCGCTGTTCCAGACCGCCCAAGGGGTCA	6884
Qy	6382	AGGTGACCAAGCGGTGTGTACCCCTGGGCCAGACATCAATGAGAGTATCACAGAGCTA	6441
Db	6885	AGGTGACCAAGCGGTGTGTACCCCTGGGCCAGACATCAATGAGAGTATCACAGAGCTA	6944
Qy	6442	CACCGGACACCAACCAAGAGCTCAGGCGACACCCCTGCGCGGCCCTTCTTACTCTTCCG	6501
Db	6945	CACCGGACACCAACCAAGAGCTCAGGCGACACCCCTGCGCGGCCCTTCTTACTCTTCCG	7004
Qy	6502	TGGGGCCAGCTGCGCCCGTCTGGAACCTTCGCGCGGCCACCCAGTGAACCTTACTTCGCG	6561
Db	7005	TGGGGCCAGCTGCGCCCGTCTGGAACCTTCGCGCGGCCACCCAGTGAACCTTACTTCGCG	7064
Qy	6562	CCCGGACCAATGTTGGCCCCCGGCGCGTGTCCCCCAACCGAAAGGGGGCAAGAGTTCG	6621
Db	7065	CCCGGACCAATGTTGGCCCCCGGCGCGTGTCCCCCAACCGAAAGGGGGCAAGAGTTCG	7124
Qy	6622	AGAGCCAAACAAACGTCGGTCTTTGGGTGTGTGAGACGGATATTGAACCTGTGTCCC	6681
Db	7125	AGAGCCAAACAAACGTCGGTCTTTGGGTGTGTGAGACGGATATTGAACCTGTGTCCC	7184
Qy	6682	ACCGAGGGGATGACCGAGGCGACAGGCGACTCCGGAAGTGTGTGACCGCTGTCGCG	6741
Db	7185	ACCGAGGGGATGACCGAGGCGACAGGCGACTCCGGAAGTGTGTGACCGCTGTCGCG	7244
Qy	6742	GGATGGGAAACAGACGAGGCCACGACGAGATGGGCTTCAAGTCTCCAGGCGAACACAGCCA	6801

Db 7245 GATGGGGAACAGACGAGCCAGCAGATGGGCTCCAGTCTCCAGGCAACAGCCCA 7304  
 Qy 6802 GCCGACGAGCTTCTTCAAGCAAGCTGACCGAGAGCAATCCGCGATGGTCAAGTCCAGAA 6861  
 Db 7305 GCCGACGAGCTTCTTCAAGCAAGCTGACCGAGAGCAATCCGCGATGGTCAAGTCCAGAA 7364  
 Qy 6862 GCAGAGATCAACAGAGCTGAAACCCCAACCGGAAATGAGCTGAAATACAAATACAG 6921  
 Db 7365 GCAGAGATCAACAGAGCTGAAACCCCAACCGGAAATGAGCTGAAATACAAATACAG 7424  
 Qy 6922 CCAGCTGGGAGCGAGATCTTCAATATGCCCCCATCACTGGAACAGGCTTATGACCTA 6981  
 Db 7425 CCAGCTGGGAGCGAGATCTTCAATATGCCCCCATCACTGGAACAGGCTTATGACCTA 7484  
 Qy 6982 TAGAAGCCAGGCGGAGTCAAGGAATGCAAGCAATGAGGCTGGAAGGCTAAATTAG 7041  
 Db 7485 TAGAAGCCAGGCGGAGTCAAGGAATGCAAGCAATGAGGCTGGAAGGCTAAATTAG 7544  
 Qy 7042 AAAGGCACTCATGGGTAAATATGACAGTGGGAAGAGTCCCGCGCTCAGCGCCAAATGC 7101  
 Db 7545 AAAGGCACTCATGGGTAAATATGACAGTGGGAAGAGTCCCGCGCTCAGCGCCAAATGC 7604  
 Qy 7102 TTTTAAACCTCTGAATGCCAGTGGCCAGCTGCTGCTATGCTATGCTCAATAAACGCTGCTGA 7161  
 Db 7605 TTTTAAACCTCTGAATGCCAGTGGCCAGCTGCTGCTATGCTCAATAAACGCTGCTGA 7664  
 Qy 7162 CGGACGAGTGAACCACTCACTCGCAGGTTGCGGAGGAGGCAAGGCTCTGCG 7221  
 Db 7665 CGGACGAGTGAACCACTCACTCGCAGGTTGCGGAGGAGGCAAGGCTCTGCG 7724  
 Qy 7222 CAGACCCAGCAGCGGAAAGCCAAAGTCCCGGCGCCGCGCTGAGCACTTGGGAGACCGGCG 7281  
 Db 7725 CAGACCCAGCAGCGGAAAGCCAAAGTCCCGGCGCCGCGCTGAGCACTTGGGAGACCGGCG 7784  
 Qy 7282 ACCCTGTCTCTCTCACTGAGCACTCGAGGAGAGCTCAACCGCGCTCAACCA 7341  
 Db 7785 ACCCTGTCTCTCTCACTGAGCACTCGAGGAGAGCTCAACCGCGCTCAACCA 7844  
 Qy 7342 CGCGGTGGGAGAGAGGCGCTCGTCCGAGGTTCAAGCCATTCGCAACCCCT 7401  
 Db 7845 CGCGGTGGGAGAGAGGCGCTCGTCCGAGGTTCAAGCCATTCGCAACCCCT 7904  
 Qy 7402 GATCATGCGGCTGCAAGCGGAGTGTCAAGCTTCCCAACCCGAGCTCCCGCGAG 7461  
 Db 7905 GATCATGCGGCTGCAAGCGGAGTGTCAAGCTTCCCAACCCGAGCTCCCGCGAG 7964  
 Qy 7462 CAGCGGCGCTCTGCGTGGCCCAACCGCTGGAAGAGAGCCCAAGCTGCTTG 7521  
 Db 7965 CAGCGGCGCTCTGCGTGGCCCAACCGCTGGAAGAGAGCCCAAGCTGCTTG 8024  
 Qy 7522 CTCGCACTGACAGCACTCTCCGACAGCGAGTCACTCAAGAGCGGCGGCGGCGG 7581  
 Db 8025 CTCGCACTGACAGCACTCTCCGACAGCGAGTCACTCAAGAGCGGCGGCGGCGG 8084  
 Qy 7582 CGGTUEBSPERLUJTCAGGTCCCAAGAGCAACAGAAACGCGCTGCAAGAGCGGAGG 7641  
 Db 8085 -----GCGGTGTCAGGTCCCAAGAGCAACAGAAACGCGCTGCAAGAGCGGAGG 8135  
 Qy 7642 GCTGCGAATCCCCCAACAGAGAGAGCCCTGAGTCCGCTGCTCATCATCT 7701  
 Db 8136 GCTGCGAATCCCCCAACAGAGAGAGCCCTGAGTCCGCTGCTCATCATCT 8195  
 Qy 7702 GTCCGTCAGAGCGCGGATCTTGGCTGTAAAGCTTAAACTCAACTCCCGCGG 7761  
 Db 8196 GTCCGTCAGAGCGCGGATCTTGGCTGTAAAGCTTAAACTCAACTCCCGCGG 8255  
 Qy 7762 CTGGCCCTGTGCAAGCTTAACTCAAGGAGATGTTTAACTGCTCGGAGAGGAGGAGAA 7821  
 Db 8256 CTGGCCCTGTGCAAGCTTAACTCAAGGAGATGTTTAACTGCTCGGAGAGGAGGAGAA 8315  
 Qy 7822 GCGGCGCGGAGAGGAGCGGAGCGGCTGTGAGGCAAGCAACAGCGCGCGAGCGGCG 7881  
 Db 8316 GCGGCGCGGAGAGGAGCGGAGCGGCTGTGAGGCAAGCAACAGCGCGCGAGCGGCG 8375

Qy 7882 AGGAGCCAAAGCAGATGACACGCACTTCAAGCCACTGCTCCCGGAAATGATTTG 7941  
 Db 8376 AGGAGCCAAAGCAGATGACACGCACTTCAAGCCACTGCTCCCGGAAATGATTTG 8435  
 Qy 7942 GAAACCAAGCTTAACTGAGCTGACAGCCCGCGGCTTCCCTGCTCCCATCCGCT 8001  
 Db 8436 GAAACCAAGCTTAACTGAGCTGACAGCCCGCGGCTTCCCTGCTCCCATCCGCT 8495  
 Qy 8002 TAGCGCTGGAGCAGATGAGACGAGGCGCTGTCAGCCCGCAGGTGGCTCCGCTC 8061  
 Db 8496 TAGCGCTGGAGCAGATGAGACGAGGCGCTGTCAGCCCGCAGGTGGCTCCGCTC 8555  
 Qy 8062 CCAGACTGCGCCAGCAGCAAGAGATGCTGGAACCAAGTCAAGCAGGAGGCGAGACA 8121  
 Db 8556 CCAGACTGCGCCAGCAGCAAGAGATGCTGGAACCAAGTCAAGCAGGAGGCGAGACA 8615  
 Qy 8122 AAAGGCGCAGGTGCGGCTGGGAGGAAAGGATGCTCGAGAGCTGATGTTTTTTCAC 8181  
 Db 8616 AAAGGCGCAGGTGCGGCTGGGAGGAAAGGATGCTCGAGAGCTGATGTTTTTTCAC 8675  
 Qy 8182 ACATGCTTCCCGCAGGCGTGGGAGGAAAGGAGATGTAATGATGTTGTTTACAG 8241  
 Db 8676 ACATGCTTCCCGCAGGCGTGGGAGGAAAGGAGATGTAATGATGTTGTTTACAG 8735  
 Qy 8242 GTATATTTTGTATCTTCAATGAATTAATTCAGATGTTTTTTCAGAAAGGACTTACC 8301  
 Db 8736 GTATATTTTGTATCTTCAATGAATTAATTAATGATGTTTTTTCAGAAAGGACTTACC 8795  
 Qy 8302 CAGTATTAATGCTGCTGCTGCTTTTGAATCTGCTTCAAGAGCGGTGACAGGCT 8361  
 Db 8796 CAGTATTAATGCTGCTGCTGCTTTTGAATCTGCTTCAAGAGCGGTGACAGGCT 8855  
 Qy 8362 GACAGTGGTGAACCCATCACTGACAGACCAAGGCGGCGGAGCTGCTGCTGCTGCTG 8420  
 Db 8856 GACAGTGGTGAACCCATCACTGACAGACCAAGGCGGCGGAGCTGCTGCTGCTGCTG 8915  
 Qy 8421 CGCTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8480  
 Db 8916 CGCTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8975  
 Qy 8481 CGCATTTGCGAGAGGTGATGATTTCTGTCATTTACACAGCTGCTTCAATTTAAAGC 8540  
 Db 8976 CGCATTTGCGAGAGGTGATGATTTCTGTCATTTACACAGCTGCTTCAATTTAAAGC 9035  
 Qy 8541 GAATTTACTCCAAAAA 8558  
 Db 9036 GAATTTACTCCAAAAA 9053

RESULT 2  
 US-09-949-016-3804  
 ; Sequence 3804, Application US/0949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; PRIOR FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3804  
 ; LENGTH: 7912  
 ; TYPE: DNA  
 ; ORGANISM: Human

US-09-949-016-3804

Query Match 7.4%; Score 630; DB 3; Length 7912;  
Best Local Similarity 65.1%; Pred. No. 4.2e-108; Indels 48; Gaps 4;  
Matches 1015; Conservative 0; Mismatches 495; Indels 48; Gaps 4;

491 CCGCCACGGCTGTCAGAGAGAGTGTATCCAGAACATGACCCGGTGGACCGAGAGATC 550  
DB CTTCAAAACTCTCAAGAGAGAGTAAATACAGAGATGATCGTGTAGATCGAGAAATTT 816  
QY 551 ACCATGTATGAGACAGACATCTCTAAGCTGAGAGAGAGAGAGAGAGAGAGAG 610  
DB GCAGAAAGTAAAGACAGACATCTTAACTGAGAGAGAGAGAGAGAGAGAGAGAG 876  
QY 611 GCTGCCAAGCCGCGAGCTGAGAGAGCCGCTGTACCGCCGCTGAGTGCAGAGC 670  
DB GCAGCTAAACCTCTGAGCTGAGAGAGCCGCTGTACCTCTCTGTGTGAGAGAGAG 936  
QY 671 CGCAGCCTGTGTCAGATCATCTACGAGAGAACCGAGAGAGAGCTGAAGCTGAC 730  
DB CGCAGTATTTGTCAGAAATTTATTTATGATGAGATCGAGAGAGAGAGAGAGAG 996  
QY 731 ATTCTGAGAGAGCTTGGGCCCCAGAGTGGAGCTGCTGTACACAGAGCTTCGAC 790  
DB 997 ATTTTGAAGGCTTGGGCCCCAGAGTGGAGCTGCTGTACACAGAGCTGACATC 1056  
QY 791 CGCAGATATCATGAGACATCAAAATTTAAACGAGGAGATCGGAGAGAGCTAT 850  
DB 1057 AAGGTATACATGAGACATCAAGACAAACGAGTGTAGAGAGAGAGAGAGAG 1116  
QY 851 TTCAAGAGAGAGAGATCACGCTCGAGAACATGAGAGAGAGAGAGAGAGAGAG 910  
DB 1117 TTTAAAGAGAGAGATCATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1176  
QY 911 CAGCTCATGAGAGAGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 970  
DB 1177 CAGCTCATGAGAGAGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1236  
QY 971 GCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1030  
DB 1237 GCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1296  
QY 1031 CGCAGCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1090  
DB 1297 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1353  
QY 1091 TCGGCGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1150  
DB 1354 ACCATTTGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1413  
QY 1151 AACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1210  
DB 1414 AATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1473  
QY 1211 CAGCAGCCGATCAAGTTCATCAACATGAGAGAGAGAGAGAGAGAGAGAGAG 1270  
DB 1474 CAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1533  
QY 1271 AAGAGCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1330  
DB 1534 AAGAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1593  
QY 1331 TTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1390  
DB 1594 TTTATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1653  
QY 1391 GCTGAGTCCGCTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1450  
DB 1654 CTTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1713  
QY 1451 AGACGAGAGATTCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1510  
DB 1714 AGAGAGAGATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1765

QY 1511 CAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1570  
DB CGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1825  
QY 1571 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1630  
DB 1826 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1878  
QY 1631 GACCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1690  
DB 1879 AATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1938  
QY 1691 GTGCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1750  
DB 1939 GCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1998  
QY 1751 CGCTCATGCTTAATGA-----GCCACAGAGAG 1780  
DB 1999 AGGTCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2058  
QY 1781 GAGGCTATCACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1840  
DB 2059 CCCCACCACTCTGCGACCGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2118  
QY 1841 CGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1900  
DB 2119 CGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2178  
QY 1901 TGTGCGGCTATCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1960  
DB 2179 TGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2238  
QY 1961 TTCACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2018  
DB 2239 TTTAATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2296

RESULT 3  
US-09-632-033B-1  
Sequence 1, Application US/09632033B  
Patent No. 6949624  
GENERAL INFORMATION:  
APPLICANT: Liu, Johnson M.  
TITLE OF INVENTION: CLONING OF THE HUMAN NUCLEAR RECEPTOR  
TITLE OF INVENTION: CO-REPRESSOR GENE  
FILE REFERENCE: NIH172.001A  
CURRENT FILING DATE: 2000-08-09  
PRIOR APPLICATION NUMBER: 60/146,977  
PRIORITY FILING DATE: 1999-08-03  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 7940  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: cDNA encoding human nuclear receptor co-Repressor  
OTHER INFORMATION: (Hun-Cor)  
US-09-632-033B-1

Query Match 7.4%; Score 630; DB 3; Length 7940;  
Best Local Similarity 65.1%; Pred. No. 4.2e-108; Indels 48; Gaps 4;  
Matches 1015; Conservative 0; Mismatches 495; Indels 48; Gaps 4;

491 CCGCCACGGCTGTCAGAGAGAGTGTATCCAGAACATGACCCGGTGGACCGAGAGATC 550  
DB CTTCAAAACTCTCAAGAGAGAGTAAATACAGAGATGATCGTGTAGATCGAGAAATTT 816  
QY 551 ACCATGTATGAGACAGACATCTCTAAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 610

```

Db      817 GCAAAAGTAAACAGCAGATCTTAAACTGAAAAAGAAACAACAGCTTGAGAAAG 876
Qy      611 GCGCCAAAGCCCGCAGCTGAGAGCCCGTGTCA CCGCCGCCATTCGATGGAAGCAC 670
Db      877 GCAAGTAAACCTCTGAGCTGAGAGCCCGTGTCTCCCTCTCTGTGAGCAAGAAACAC 936
Qy      671 CCGAGCTGTGTGAGATCTCTACGACGAAACCGGAGAGAGAGCTGACATCGG 730
Db      937 CGCAGATATGTCCAATTTATTTATGATGAGAAATCGGAAAAAGCAAGAAAGCTCATAAA 996
Qy      731 ATTTGAAAGGCTTGAGGAGCCCAAGGTGAGCTGC CGCTGTACAACAGCCCTCGACCC 790
Db      997 ATTTTGAAGGTCTTGAGCCCAAAAGTTGAACTGCCA CTGTATTAACAGCCATCAGATACC 1056
Qy      791 CCGAGATCATAGAAACATCAAAATTAACAGAGCAATGCGGAGAGAGCTTAATCTTGAC 850
Db      1057 AAGGTACCATAGAAACATCAAGAACAAACAGGTGATGAGAAAAAACTCATTTATTT 1116
Qy      851 TTCAAGAGAGAGAAATCACTCGGAAACAAATGGAACAGAAATTCTGCCAGCTATGAC 910
Db      1117 TTTAAAGAAAGAAATCATGCAAGAAACAAAGGAAACAAAAATCTGCCAGCTTATGAT 1176
Qy      911 CAGCTCATGAGGCTTGAAAAAAAGGTGAGCGCATCGAAAAACAACCGCGCGCGCG 970
Db      1177 CAGCTCATGAGGCAATGAGGAAAAAAGTGAACAGATAGAAAAATATCTCTCGAGGAAA 1236
Qy      971 GCCAAGAAAGCAAGGTGCGCGAGTACTACGAAACAGATCTCTGATTCGCAAGCAG 1030
Db      1237 GCTAAAGAAAGCAAAACAGGGAATCTATGAAACAGATTCACAAATTCGAAAAACA 1296
Qy      1031 CCGAGCTCGAGAGGCAATGACAGACAGAGGTGAGCGCGAGCGGAGCTGTCCATG 1090
Db      1297 AAGAAACAGACAAAGAAATTTGAG---CGAGTTGGGAGAGGAGACTGCTTTAGACC 1353
Qy      1091 TCGGCGCGCCGAGCGAGCAAGAGGTGTCAAGATCATGATGCGCTTCAGACAGAG 1150
Db      1354 ACCATGTGATAGAGGTAGAGATGATTTCTGAAATTAATGATGGGCTCTCTGAGCAGAG 1413
Qy      1151 AACCTGGAAGACAGATGCGCCAGCTGAGCTGATCTCCGCCATGCTGTACAGACGTGAC 1210
Db      1414 AATTAATGAACAAATGCGGCGCTCTGTGATTCACATGATGATTTGATGAGAA 1473
Qy      1211 CAGCAGCGCATCAAGTTCAATCAATGAACGAGCTTATGAGCGAGCCCAATGAAGTGTAC 1270
Db      1474 CAAGAAGAGTCAAGTTCAATTAATGAATGAGCTTATGAGAGACCTTATGAAGTGTAT 1533
Qy      1271 AAAGACCGCAGGTCAATGATGTGAGTGAACAGAGAGAGAGAGAGAGAGAGAGAG 1330
Db      1534 AAAGATAGGCAATTTATGATGATTTGAGCTGACCATGAAGAGAGATCTTTAAGAGAGAG 1593
Qy      1331 TTCAATGACATCTCCAAAGAACTTTGGCTGTATGCTCATTTCTGAGAGAGAGAGAGAG 1390
Db      1594 TTTATTCAGACATCAAAAACTTTGAGCTTAATTCATTAATTCGAGAGAGAGAGTGT 1653
Qy      1391 GCTGAGTGTCTCTCAATTAATACTGATGAAGAAATGAAGAACTATTAAGAGCTGTG 1450
Db      1654 CTTGATGTGTTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1713
Qy      1451 AGACGAGCTATTCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1510
Db      1714 AGAAGGAATTAATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1765
Qy      1511 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1570
Db      1766 CCGAAGAAAGAAAGTGAAGAAAGAAAGAGAGATTAAGCAAGAAAGAAAGAAAGAAAG 1825
Qy      1571 GAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1630
Db      1826 AAGAGAGAAAG-----GAAAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1878
Qy      1631 GACCTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1690
Db      1879 AATTAACAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1938

```

```

Qy      1691 GTGCTCTCCAAAGGCGCGCAAAATCTGCCAAACGCCAGAGGAAAGAGCAAAAGGCGCATCAC 1750
Db      1939 GCCACACCCCGGAGGAGGAAAGACTGCCAACAGTCAAGGGCCGCGCTTAAGGGCCGATCAC 1998
Qy      1751 CGCTCAATGCTAATGA-----GGCCAAACAGCAG 1780
Db      1999 AGGTCCATGACAAACGAAGCTGACGTGCGAAGTGTGACAGCGCAGAGGCTTATGAAAGAG 2058
Qy      1781 GAGGCCATCACCCCCAGCAGAGCGCGAGCTGCTCCATGAGAGCTGAATGAGATTCT 1840
Db      2059 CCCCCCAACCTCTGACACCGCCACCAAGAACCCATTTCTACAGAGCCTGTGAGAGACTCT 2118
Qy      1841 CGCTGACAGAGAAAGAAATGAGAAACAGCCAGAAAGGTCTCTGAGAAACGCGCGCAAC 1900
Db      2119 CGATGACAGAGAAAGAAATGAGAAAGTGTCTTAAAAAGGTCTAGTAAACATGATCTTAAC 2178
Qy      1901 TGGTGGCCATCGCCCGAGTGGTGGCTCAAGACTGTGTGCAATGTAAAGACTTCTAC 1960
Db      2179 TGGGACGCAATTGCTAAATGGTGGGACGAAAGTGAAGCTCAATGTAAACCTCTAT 2238
Qy      1961 TTCACTAACAAGAGAGGAGCAACCTCGATGAGATTTGACAGAGACAAAGCTGAAGA 2018
Db      2239 TTTAACTATAAAGGCGACACATCTTGAACAACCTCTTACAGCAGCATAAACAGAAA 2296

RESULT 4
US-09-632-033B-2
; Sequence 2, Application US/09632033B
; Patent No. 6949624
; GENERAL INFORMATION:
; APPLICANT: Liu, Johnson M.
; APPLICANT: Wang, Jianxiang
; TITLE OF INVENTION: CLONING OF THE HUMAN NUCLEAR RECEPTOR
; FILE REFERENCE: NIH172,001A
; CURRENT APPLICATION NUMBER: US/09/632,033B
; CURRENT FILING DATE: 2000-08-09
; PRIOR APPLICATION NUMBER: 60/146,977
; PRIOR FILING DATE: 1999-08-03
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 7780
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: cDNA encoding murine nuclear receptor co-receptor
; NAME/KEY: misc_feature
; LOCATION: (1)...(7780)
; OTHER INFORMATION: n = A,T,C or G
US-09-632-033B-2

Query Match      7.4%; Score 629.6; DB 3; Length 7780;
Best Local Similarity 64.4%; Pred. No. 4,9e-108;
Matches 1029; Conservative 0; Mismatches 524; Indels 45; Gaps 4;

Qy      446 TCTCCCCCAGGCCCCCGGACACTGACCTGAGCTGAGACTGCTGTCGCGCAGGCTGCC 505
Db      588 TCTCCCTCTCTGGGAGGCGCATGCGAGATGATCAAGATGCTCAACCTTCAAAACGTCA 647
Qy      506 AAGAGAGAGCTATTCAGAAACATGACCGCGTGAACCGAGAGATCAACATGTGAGCAG 565
Db      648 AAGGAAGAGCTATTCAGAGCATGATCGTGTGATCGAGAAATGCGAAAGTGAAGACAG 707
Qy      566 CAGATCTCTAAGCTGAAGAAAGAGAGCAACAGCTGGAAGAGAGAGGCTGCAAGCGGCC 625
Db      708 CAGATCTTAAACTGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 767
Qy      626 GAGCTGAGAGAGCCGCTGTCAACCGCGCCCATCGATCGATCGAAGCAACGAGCCTGTGAG 685

```



Db 768 GAGCTGAGAAAGCTGTGTCCCTCCCTCCCTGAGAGCAGAAAGCAGGAATATGTTCCAA 827  
Qy 686 ATCATCTACGACGAGAAACCGGAGAAAGGCTGAAGCTGCAATCGGATTTCTGGAAGGCTTG 745  
Db 828 ATCATTTATGACGAGAAATCGGAAAAAGAGAGAAAGCTCAATAAATATTTGAAGTCTT 887  
Qy 746 GGGCCCCAGGTGAGCTGCGCTGTACAAACAGCCCTTCGACACCCGGCAGATATCATGAG 805  
Db 888 GGGCCCAAAAGTTGAACCTGCGCTTACAAACGCGGTGAGATACCAAGGTGATACACAG 947  
Qy 806 AACATCAAAATTAACGAGCCGATGCGGAAAGAACTAATTTGTACTTCAAGAGAGAAAT 865  
Db 948 AACATCAAGACAAACAGGTGATGAGAAAAAATCTCATTTATTTTAAAAAGAAAT 1007  
Qy 866 CACGCTCGGAAACAATGAGCAGAGAAATTTCTGCAAGCTATACAGCTCATGAGGCGC 925  
Db 1008 CATGCAAGAAAAACAAGGGAACAAAAATCTGCAACGTTATGATCAGTCTATGGAAGCA 1067  
Qy 926 TTGAAAAAAAGGTGAGCGCATCGAAAAACAACCGCGCGCGGCGCCAGAGAGACAG 985  
Db 1068 TGGGAGAAAAAGTGGACAGATAGAAAAATATCTCGGAGGAAAGCTTAAAGAAACAAA 1127  
Qy 986 GTGGCGAGTACTAGAAAAAGAGTTCCCTGAGATCCGCAAGCAGCGAGCTGAGAGAG 1045  
Db 1128 ACAGAGGAATACTATGAAAAAGCAGTTTCCAGAAATTCGAAAAACAAGAGACAGAGAA 1187  
Qy 1046 CGCATGACAGAGCAGGAGTGGCCAGCGGCGAGTGGGCTGTCCATGTGCGCCGCCAGC 1105  
Db 1188 AGATTTTCAG---CGAGTTGTGTAGAGGGAGGAGTGGTCTTTTCAGCCACCTTGTGAGAGT 1244  
Qy 1106 GAGCAGAGGTGTCAAGATCATCGATGCGCTCTCAAGACAGAGAAACCTGGAAGACAG 1165  
Db 1245 GAGCATGAGATTTCTGAATTTATGATGCTTTCTGAAACAGAGAAATATGAGAAAGCA 1304  
Qy 1166 ATGGCGCAGCTGGCGGTATCCCGCCCATGCTGTACAGCGTCAAGAGGAGCATCAAG 1225  
Db 1305 ATGCGTCACTTTCGTGATTTCCACTTATGATGTTTGAATGCAAAACAAGAGGTCAAA 1364  
Qy 1226 TTTCATCAACATGAACGGGCTTATGGCCGACCCCATGAAGGTGTACAAAAGCCGACAGTC 1285  
Db 1365 TTTCATCAATATGAATGGGCTGATGAGAGATCCAAATGAAGTTTATTAAGACAGACAGTTT 1424  
Qy 1286 ATGAACATGTGAGTGAAGCAGAGAAAGAGACCTTCCGGAGAGGTTCAATGACATCCC 1345  
Db 1425 ATGAATGTTTGGACATGACATGAAAAAGAGATCTTTAAGACAAAGTTTATCCAGATCCA 1484  
Qy 1346 AAGAACTTTGGCTGTATCGCATCATTTCTGAGAGAGAAAGACAGTGGCTGAGCGTCTCC 1405  
Db 1485 AAAAACTTTGACCTAATTTGCATCTTATTTGAAAGGAAAGTGTTCCTGATTTGTGTTTA 1544  
Qy 1406 TATTACTACTGACTTAAGAAATGAAGAACTATTAAGGCTGTGTGAGACGAGCTATCCG 1465  
Db 1545 TATTTCTATTTTAACCAAGAAAAATGAATTTAATAGCCCTGTGAGAGAAATTAATGGA 1604  
Qy 1466 CGCCGCGCAGAGGCCAGCAGCAACAACGACGACGACGACGACGACGACGACGACGAC 1525  
Db 1605 AAACGCAAGGCGAGAAATCAGCA-----GATTCGCGTCCCTCAACAAGAG 1650  
Qy 1526 CAGCAGCCCATTCGCCCGCAGCAGCCAGAGAGAAAGATGAGAGAGAGAGAGAGAGAG 1585  
Db 1651 AAAAAAGTAAAGAAAAAGAAAGAGATTAAGCAGAAAAACAG-AGAAAAAGAGAAAGAA 1709  
Qy 1586 GCGGAGAGAGAGAGAGAGAGCCGAGAGGTGAGAAACGACAGAGAAAGACCTTCTCAAGAG 1645  
Db 1710 AAGAAAGATGATGAAGAAAAAATATGATTAAGAGAGACTTAAGAGAAACAACAGAGAAAG 1769  
Qy 1646 AAGACAGACGACCTCAGGGAGAGCAACGACGAGAGAGAGCTGTGGCTTCCAAAGGC 1705  
Db 1770 GACAGAGCGAAGCAGCAGCAGAGAAAGAACTGAAAGAAAGAGACAGGTCACTCCAAGGGG 1829  
Qy 1706 CGCAAAATTCGCACAGCCAGGAGAGAGCAGAAAGCCGATCATCCGCTCAATGGCTAAT 1765  
Db 1830 CGAAAGACTGTAAACAGCCAGCGCGGGAGAGGCGGGGTCAACAGGTGATGACAACT 1889

Qy 1766 GAG-----GCACAAGCAGAGAGGCCATCAACCCCCAG 1798  
Db 1890 GAAGCTGACGCTGCCAATGTCTGTCTGCAAGCTCAGTGAAGAGCCGCCCAACCCCTGCGG 1949  
Qy 1799 CAGAGCCCGAGCTGCGCTCCATGAGAGCTGAATGAGATTTCTGCTGGAACAGAAAGAA 1858  
Db 1950 CCACCACAGAGCCCATTTCTACAGAACTCTTGAAGATTTGAAGATGAGACAGAAAGAA 2009  
Qy 1859 ATGGAACAGCCAAAGAAAGTCTCCTGGAACAGCGCCGCAACTGCTGCGCCATCGCCCG 1918  
Db 2010 ATGGAAGTCTTAATAAAAGCCCTGATAGACATGTGCTTAATCTGGCAGCCATTTCTAAA 2069  
Qy 1919 ATGTGGGCTCCAGACATGTGTGCGAGTGAAGAACTTTCTAATTCACTACACAGAGAG 1978  
Db 2070 ATGGTGGGAACTAAGAAAGTGAAGCCAGTGAACAACTTCTATTTTAATATTAAGACGG 2129  
Qy 1979 CAGAACTTCGATGAGATTTTGCAGAGACAGCAAGCTGAA 2016  
Db 2130 CATATCTTGAACACCTTTTGCAGCAACATTAACAGAA 2167

RESULT 5  
US-09-513-999C-33415  
; Sequence 33415, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Ducleit, A.Y.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59, US2, REG  
; CURRENT APPLICATION NUMBER: US/09/513, 999C  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent, pm  
; SEQ ID NO 33415  
; LENGTH: 312  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 86  
; OTHER INFORMATION: s=g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 87  
; OTHER INFORMATION: m=a or c  
US-09-513-999C-33415

Query Match 3.6%; Score 311.2; DB 3; Length 312;  
Best Local Similarity 99.4%; Pred. No. 6.3e-49;  
Matches 310; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1560 AAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1619  
Db 1 AAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60  
Qy 1620 ACGACAGAGAACTCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1679  
Db 61 ACGACAGAGAAAGCTCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
Qy 1680 AGAGAGAGCTGTGCTCTCCAAAGGCGCGCAAACTGCGCAACGCTCAGAGAGAGAGAGAG 1739  
Db 121 AGAGAGAGCTGTGCTCTCCAAAGGCGCGCAAACTGCGCAACGCTCAGAGAGAGAGAGAG 180  
Qy 1740 GCCGATCAACCGCTCAATGGCTAATGAGGCCAACAGGAGAGAGGCTATCACCCCTCAGC 1799  
Db 181 GCCGATCAACCGCTCAATGGCTAATGAGGCCAACAGGAGAGAGGCTATCACCCCTCAGC 240

QY	1800	AGAGGGCCAGCTGGCCCTTCATGAGCTGATGAGATGAGATTCCTCGCTGACAGAAAGAAA	185
Db	241	AGAGCGCCGAGCTGGCCCTTCATGAGCTGATGAGATTCCTCGCTGACAGAAAGAAA	300
QY	1860	TGAAAACAGCCA	1871
Db	301	TGAAAACAGCCA	312
RESULT 6			
US-08-372-652-9			
; Sequence 9, Application US/08372652			
Patent No. 5932699			
GENERAL INFORMATION:			
APPLICANT: Moore, David			
APPLICANT: Seol, Wongi			
APPLICANT: Choi, Hwang-Sik			
TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING			
TITLE OF INVENTION: POLYPEPTIDES AND RELATED MOLECULES AND METHODS			
NUMBER OF SEQUENCES: 17			
CORRESPONDENCE ADDRESS:			
ADDRESS: Fish & Richardson P.C.			
STREET: 225 Franklin Street, Suite 3100			
CITY: Boston			
STATE: MA			
COUNTRY: USA			
ZIP: 02110-2804			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/372,652			
FILING DATE: 13-JAN-1995			
ATTORNEY/AGENT INFORMATION:			
NAME: Clark, Paul T.			
REGISTRATION NUMBER: 30,162			
REFERENCE/DOCKET NUMBER: 00786/246001			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: 617/542-5070			
TELEFAX: 617/542-8906			
TELEX: 200154			
INFORMATION FOR SEQ ID NO: 9:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 1922 base pairs			
TYPE: nucleic acid			
STRANDEDNESS: single			
TOPOLOGY: linear			
MOLECULE TYPE: DNA			
US-08-372-652-9			
Query Match 1.7%; Score 142; DB 2; Length 1922;			
Best Local Similarity 52.5%; Pred. No. 4.4e-17;			
Matches 435; Conservative 0; Mismatches 375; Indels 18; Gaps 5;			
QY	6730	GCTGCTGTACCGGAGTGGGAAACAGACGGAAGCCGACGAGATGGGCTTCCAGTCTCCAGG	6789
Db	1039	GCTTGTTCACAGAGGGAGAGTGAGACCTCGTGGAGCAAGAGATGATTTCTGATACCAAGG	1098
QY	6790	CAACACACGACGACCGCGCCAGCTTCTTCAGCAAGCTGACCGAGAGCAATCCGCCATGGT	6849
Db	1099	AAGTATTAAGCTACTTGGCTTCAATCTTCCACCAAGCT--TGAAAGCATATCAACCAATGGT	1155
QY	6850	CAATTCAGAGAGGAAAGATCAACAGAAAGCTGAACCCACAAACCGGAATGAGCTTGA	6909
Db	1156	TAAATCAAGAAACAGGAAATTTTTTCTGAATGAACCTTCTGGTGAAGGTGACTTGA	1215
QY	6910	ATACAAATACGACGACCTTGGAGCGAGAGATCTTCAATATGCCCCCATCAACGGAACAGG	6969
Db	1216	TATGGCAGCTGCTCAGCCAGGAACAGAGATCTTCAATGTGCAAGACGATTAACAATCAGG	1275
QY	6970	CCTTATGACCTATATGAAAGCCAGGCGGTGACAGAAATGCCAGACCAATATGGGGCTGGA	7029

```

Db      1276  TGCAGTAGGCTCAGAAAGCCATTCTTTTGTGCGATCCGCCAG---TAACTTGGTCTAGA 1332
Oy      7030  GGCCATAATTAAGAAAGCATCTATGGGTAAT---ATGACCATGGGGAGAGTCCCCGCC 7086
Db      1333  AGACATCATCAGAAAGGCTCTCATGGGAAAGTTTGAATGATTAAGTTGAAAGTCAATGTGT 1392
Oy      7087  GCTCAGCGCCCAATGCTTTTAAACCTCTGAAATGCCAGTGGCCAGCCCTGCCCGCTGATGCC 7146
Db      1393  TGTATATGTCCTATCTGTGGGCAATTATGCTCGTATGAGCCAGCACCTTATGATGAGAG 1452
Oy      7147  CATTAACCGCTGCTGACGAGACGGAGTGAACAACAACCTACCTCGCAGGTGGCGCGGAA 7206
Db      1453  CAGGAGAGGACCGAGAGATGAAGGGAGCCATCACTCATGCAAGAGATATGCAAACCAA 1512
Oy      7207  GGCCAAAGTCTTGTGGAGAACCCAGACAGCCGAAAAGCCAAAGTCCCCCGGCGCTGGC 7266
Db      1513  GCTGATACAAATCAAAACAGCAGAGAGTCAATATCTCTATTCCTGGGGCAAGACTATTT 1572
Oy      7267  ATCTGGGAGACGGGACACCTCTGTCTCTCTCATGTGCACTCGGAGGAGATGCAACCGCG 7326
Db      1573  AGGAATGAAAGGCGCTTCTGTCTGTCTCTCTGTGCAATTCAGAAAGTATTAACCAAGCA 1632
Oy      7327  GAGCGCGCTCAACAACCGCGGTGGGAGGACAGGCCCTCGTCCGACAGTTCACGCCATT 7386
Db      1633  GACAC-----CAGATGGGATGGAAAGATCGGCCCTTTCAACAGAGTCTACTCAATT 1686
Oy      7387  CCCCTTAACACCCCTGATCAATGCGGCTGCAAGGCGGATGTCATGGCTTCCCAACCCGAC 7446
Db      1687  CCTTAAACACCTCTGACATACATAGCGATGCTCAGCAGTACACACACTACAGAGATGGATG 1746
Oy      7447  GAGCTCTCCCGCGGGGAGC---GGGCCCCCTGCTGCGCCCCCAACGCGCTGGGAGAGGA 7503
Db      1747  CGCCCATCTGCGCATCAACCCAGACGCTCCACATCAACAGAACCGCATCTGGGAGAGGGA 1806
Oy      7504  GCCCAACCACTGCTGCTGCGAGTACGAGACCTCTCCGACAGGA 7551
Db      1807  GCCGCGCCGCTCTCTCAGCGCAGTATGAGACACTGTGATATGTA 1854

RESULT 7
PCT-US95-16311-9
Sequence 9, Application PC/TUS9516311
GENERAL INFORMATION:
APPLICANT: Moore, David
APPLICANT: Seol, Wong
APPLICANT: Choi, Hwang-Sik
TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING
TITLE OF INVENTION: POLYPEPTIDES AND RELATED MOLECULES AND METHODS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street, Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16311
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/372,652
FILING DATE: 13-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/246001
TELECOMMUNICATION INFORMATION:

```

TELEPHONE: 617/542-5070  
 TELEFAX: 617/542-8906  
 TELEEX: 200154  
 INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1922 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA  
 PCT-US95-16311-9

Query Match 1.7%; Score 142; DB 6; Length 1922;  
 Best Local Similarity 52.5%; Pred. No. 4.4e-17;  
 Matches 435; Conservative 0; Mismatches 375; Indels 18; Gaps 5;

6730 GCTGCTGACCGGGATGGGGAACAGACGAGCCGACGAGGATGGGCTCCAACTCCACG 6789  
 1039 GCTCTTGTCACAGAGGGAGGTGGACCTGTGAGCAAGAGATGATTCGATCCACAGG 1098  
 6790 CAACACGACCGCCGACGACCTCTTTCAGCAAGCTGACCGAGCAACTCCGCAATGGT 6849  
 1099 AAGTATAGCTACTTGCTTCAATCTTTCACCAAGCT---TGAAGCAGATCACCATGGT 1155  
 6850 CAAGTCCAGAAAGCAAGATCAACAGAGCTGAAACCCCAACCGGATGAGCTTGA 6909  
 1156 TAAATCAAGAAACGAAATTTTCTGTAAGTTGAATCTTCTGGTGGAGGTGACTTGA 1215  
 6910 ATACAAATTCAGCGCCGCTGGAGCGGAGATCTTCAATATGCCCCGATACCGGAACGG 6969  
 1216 TATGGCAGCTGCTCAGCGAGCAAGAGATCTTCAATCTGCAAGAGATTAACATCAGG 1275  
 6970 CTTATGACTATATAGACGACGCGGTGCGAGACATGACCAACCAATGAGGCTGGA 7029  
 1276 TGCACTGAGCTCAAGAAAGCATCTTTTGCTGATCCCGCAG---TAACTTGGCTAGA 1332  
 7030 GGCATTAATGAAAGCACTCATAGGTAAT--ATGACAGTGGGAAAGATCCCCGCC 7086  
 1333 AGACATCATCAGAAAGGCTCATAGGAAAGTTTATGATTAAGTTAAAGATCATGTGT 1392  
 7087 GCTCAGCGCCAAATGCTTTTAACCCCTGTAATGCCAGTGCACGCTGCCCGCTGATGCC 7146  
 1393 TGTCTATGTCCTCATCTGTGGGATTAATGCTGTGATGTCACACCTCAGTGTACGAG 1452  
 7147 CATTAACCGCTGTGACGACGAGATGACACACACTCACTGCGCAGGTGGCGGGGAA 7206  
 1453 CAGCGAGGCAAGGAGATGAAAGGGAGCATCACTCATGAGAGATGAAACCAA 1512  
 7207 GACCAAGGCTCTTGGCAGACCCAGACGCGAAAGCCAAATCCCGCGCCCGGCTGGC 7266  
 1513 GCTGATCAACAAATCAACAGAGAAAGCTAAATCTCTATTCTGGGCAAAAGCTATT 1572  
 7267 ATCTGGGAGACCGGCAACCTCTGTCTCTCTCACTGACCTCGGAGGAGACTGAAACCGCG 7326  
 1573 AGGAATCAAAAGGCTTCTTCTGTCTCTCTGTGATTCAGAAAGGTGATTAACAGAGGA 1632  
 7327 GACGCGCTCAACCAACCGGTGAGGAGGACAGGCGCTGTCGAGAGGTTCACGCCATT 7386  
 1633 GACAC-----CAGGATGGGATGGAAGATGGGCGCTTCAACAGGTTTCACTCAATT 1686  
 7387 CCCCTCAACCCCTGATCATGCGGCTGACGCGGGGTGTCAATGAGTTCCCAACCCCAAC 7446  
 1687 CCTTCAACACCTTCAACATACGATGCTGACAGTACACCACTTACCAATGCAATG 1746  
 7447 GGGGCTCCCGGGGAGC---GGGCGCTGCTGGCGCCCAACAGCTTGGAGAGGA 7503  
 1747 GCGCCCATCTGCAATACCAAGAGCTTCAATCAACGAACCCGATCTGGAGAGGA 1806  
 7504 GCCCAAGCACTGCTGCTGCTGCAATGACAGACACTTCCGACAGCA 7551  
 1807 GCTGCGCCGCTCTCTCTCAGCGCAGTATGAGACACTGCTGTGATGTA 1854

RESULT 8  
 US-08-728-323A-1  
 Sequence 1, Application US/08728323A  
 Patent No. 5948676

GENERAL INFORMATION:  
 APPLICANT: Chang, Yuan  
 APPLICANT: Bohenzky, Roy A.  
 APPLICANT: Russo, James J.  
 APPLICANT: Edelman, Isidore S.  
 APPLICANT: Moore, Patrick S.  
 TITLE OF INVENTION: Immediate Early Protein From Kaposi's  
 TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA  
 TITLE OF INVENTION: Encoding Same And Uses Thereof  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Cooper & Dunham LLP  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/728.323A  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: White, John P.  
 REGISTRATION NUMBER: 28,678  
 REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-278-0400  
 TELEFAX: 212-391-0525  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3489 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..3489  
 US-08-728-323A-1

Query Match 1.3%; Score 107.8; DB 2; Length 3489;  
 Best Local Similarity 43.8%; Pred. No. 1.3e-10;  
 Matches 631; Conservative 0; Mismatches 792; Indels 18; Gaps 3;

219 AGCCGGGAATGAACGGTCCAGAGCTTCACCTGCGGCCAGAGTCCACTCATACCTGC 278  
 1352 AGCCACAGCAGAGAGGCAAGCAGCAGAGAGCCCTTCAGAGAGCCAAACAGCAGAGC 1411  
 279 CCGAGCTGGGGAAGTCAGAGATGAGTTCATTGAAGAGAGCGCCCTGCGCTAGAGCTGC 338  
 1412 CACAGCAGAGAGGCAAGCAGCAGAGAGCCCTTCAGAGAGCCAAACAGCAGAGC 1471  
 339 TGCCCTGACCCCTGCTGCAACCTGTCACCTGCTGCGCCACAGGAGCCTGCGGATCTG 398  
 1472 AGCAGCAGAGAGCCCTTGAAGAGGCAACAGCAGAGAGCCAAACAGCAGAGC 1531  
 399 AAGACTTACCAAGAGCCGTAGCTGACGCGGAGAGCTGAAACCGGTGTCTCCCCCAAGC 458  
 1532 AGCAGAGCCACAGCA-----GCAGAGCACAAGCAGCAGAGAGCACAAGCAGCAGAGC 1585  
 459 CCCGCAACACTGACCTGAGCTGAGCTGTGCGCGCCAGCGCTGTCCAAGAGAGAGCTGA 518  
 1586 CACAGCAGAGAGCCACAGCAGCAGAGAGCCACAGCAGAGAGCACAAGCAGCAGAGAGC 1645

Qy 519 TCCAGAAATGGAACCGCTGTGACCGAGAGATCACTATGTTAGACAGAGATCTTAAGC 578  
Db 1646 CACAGACGCGGGAGCCACAGCAGCGGGAGGCCACGACGCGGGAGCCACAGCAGCGGGAGC 1705  
Qy 579 TGAAGAAGAGCAGCAGCTGTGAGAGAGGCTCCAAAGCCGCCAGCTTGAGAGC 638  
Db 1706 CACAGACGCGGGAGCCACAGCAGCGGGAGCCACAGCAGCGGGAGCCACAGCAGCGGGAGC 1765  
Qy 639 CCGTGTCAACCGCGCCCATCTGAGTCCAGAGCAGCGCTGTGAGATCATCTAGAGC 698  
Db 1766 CACAGACGCGGGAGCCACAGCAGCGGGAGTGAAGCAGCAGGATGAGCAGCAGCAGAGATG 1825  
Qy 699 AGAACCAGAAAGAGCTGAGCTGCACTCGAATCTTGAAAGCCTGAGGCCCCAGAGTG 758  
Db 1826 AGCAGCAGCAGGATGAGCAGCAGCAGAGATGAGCAGCAGGATGAGCAGCAGCAGAGATG 1885  
Qy 759 AGCTGCGCTGTACACAGCAGCCTCCGACACCCGAGATGTCACTGAGAAATCAAAATTA 818  
Db 1886 AGCAGCAGCAGGATGAGCAGCAGCAGAGATGAGCAGCAGCAGGATGAGCAGCAGAGATG 1945  
Qy 819 ACCAGCGATGGGAGAGAGTAACTTGTACTTCAAGAGAGAGATTCAGCGCTGGAGAC 878  
Db 1946 AGCAGCAGCAGGATGAGCAGCAGCAGAGATGAGCAGCAGCAGGATGAGCAGCAGAGATG 2005  
Qy 879 AATGGAAGCAGAAATTCCTGCAAGCGCTATGACCAAGCTCATGAGAGCCTTGAAAAAAG 938  
Db 2006 AGCAGCAGCAGGATGAGCAGCAGCAGAGATGAGCAGCAGCAGGATGAGCAGCAGAGATG 2065  
Qy 939 TGGAGCGCATCGAAAACAACCGCGCGCGCGCGCCAAAGAGAGCAAGTGGCGAGTACT 998  
Db 2066 AGCAGCAGCAGGATGAGCAGCAGCAGAGATGAGCAGCAGCAGGATGAGCAGCAGAGATG 2125  
Qy 999 AGGAAAGAGATTCCTGTGATTCGGAACAGCAGCGAGCTGACAGGCGCGATCAGAGCA 1058  
Db 2126 AGGATGAGAG-----GAGCAGCAGAGATGAGCAGCAGGATGAGCAGCAGCAGCAGC 2179  
Qy 1059 GGGTGGGCGAGCGGGGCACTGGGCTGTCCATGTCCGCGCCGCAAGCAGCAGAGTGT 1118  
Db 2180 ATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGCAGCAGAGATGAGCAGCAGCAGAGATG 2239  
Qy 1119 CAGAGATCATCGATGCGCTCTCAGAGCAGAGAACTTGAGAAAGCAGATGCGCAGCTGG 1178  
Db 2240 AGCAGCAGCAGATGAAACAGAGCAGCAGAGAGAGAGAGAGCAGAGCAGAGCAGAGC 2299  
Qy 1179 CCGTGTATCCCGCCCATGCTGTACAGAGCTGACAGCAGCGCATCAAGTTTCAATAATGA 1238  
Db 2300 AGGATTTAGAGAGCAGAGCAGAGAGATTTAGAGATCAGAGCAGAGATTTAGAGAGCAG 2359  
Qy 1239 AGCGCTTATGGCCGACCCCATGAGATGTAAACAAAGACCGCCAGTCAATGAACATGTGA 1298  
Db 2360 AGCAGAGATTTAGAGAGCAGAGCAGAGAGATTTAGAGAGCAGAGCAGAGATTTAGAGAGC 2419  
Qy 1299 GTGAGCAGAGAGAGAGACCTTCCGGAGAAATTCAATGACATCCCAAGAACTTTG--- 1355  
Db 2420 AGGAGCAGAGATTTAGAGAGCAGAGCAGAGAGATTTAGAGAGCAGAGCAGAGATTTAGAG 2479  
Qy 1356 ---GCCGTATCCCATATCTCTGTGAGAGAGAAAGCAGTGGCTGAGTCCCTCTATTACT 1412  
Db 2480 AGCAGAGAGAGAGATTTAGAGAGCAGAGAGATTTAGAGAGCAGAGAGTGAAGAGC 2539  
Qy 1413 ACTTGACTTAAGAAATGAGAACTTAAGAGCTGTGTGAGACGGAGCTATCGCGCGCGG 1472  
Db 2540 AAGAGCAGAGAGTGTGAAGAGCAGAGCAGAGAGAGAGAGAGAGAGATTTAGAGAGAGTGG 2599  
Qy 1473 GCAAGAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1532  
Db 2600 AGGAGCAGAGAGCAGAGCAGAGAGCAGAGAGAGAGAGAGAGATTTAGAGAGTGAAGAGC 2659  
Qy 1533 CCATGCCCCCGCAGCAGCAGCAGAGAGAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1592  
Db 2660 AGGAGAGAGAGAGATTTAGAGAGAGTGAAGAGAGAGAGAGAGAGATTTAGAGAGAGTGG 2719  
Qy 1593 AGGAGAGAGAGAGAGCGAGAGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1652

Db 2720 AAGAGCAGAGCAGCAGAGATTTAGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 2779  
Qy 1653 A 1653  
Db 2780 A 2780  
  
RESULT 9  
US-09-298-568-1  
; Sequence 1, Application US/09298568  
; Patent No. 6322792  
; GENERAL INFORMATION:  
; APPLICANT: Kieff, Elliott D.  
; APPLICANT: Ballester, Mary E.  
; APPLICANT: Kave, Kenneth M.  
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO  
; FILE REFERENCE: 16412-10001R  
; CURRENT APPLICATION NUMBER: US/09/298,568  
; EARLIER FILING DATE: 1999-04-21  
; EARLIER APPLICATION NUMBER: US 60/109,422  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3489  
; TYPE: DNA  
; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-298-568-1  
  
Query Match 1.3%; Score 107.8; DB 3; Length 3489;  
Best Local Similarity 43.8%; Pred. No. 1.3e-10;  
Matches 631; Conservative 0; Mismatches 792; Indels 18; Gaps 3;  
  
Qy 219 AGCCCGGAGATGAAACGCTCCAGAGCTCCAGAGCTCGAGGCAAGTCCACTATACCTGC 278  
Db 1352 AGCCACAGCAGAGAGAGCCACAGCAGCAGAGAGCCCTGACAGAGGCCAACAGCAGAGC 1411  
Qy 279 CCGAGCTGGGAGAGTGAAGATGAGATTCATTGAAAGCAAGCGCCCTCGCTGAGACTGC 338  
Db 1412 CACAGCAGCAGAGAGCCACAGCAGCAGAGAGCCCTGACAGAGGCCAACAGCAGAGAGCCAG 1471  
Qy 339 TGCCTGACCCCTGCTGCGACCGCTCACCCCTGTGCGCAAGGCGCAGCTGCGAGATTCG 398  
Db 1472 AGCAGCAGAGAGCCCTGCTGAGAGCCACACAGAGAGCCACACAGCAGAGAGCCAGC 1531  
Qy 399 AAGACTTCAACAAAGACCGTACCTGACCGGCAAGCTGGAACCGATGTCTCCGCCAGCC 458  
Db 1532 AGCAGAGCCACAGCA-----GCAGAGCCACAGCAGAGAGGCCACAGCAGCAGAGC 1585  
Qy 459 CCCCAGCATGTAACCTGAGCTGAGAGCTGTGCGCCAGCTGTCTCAAGAGAGAGCTGA 518  
Db 1586 CACAGCAGCAGAGCCACAGCAGCAGAGAGCCACAGCAGAGAGAGCCACAGCAGAGAGC 1645  
Qy 519 TCCAGAAATGGAACCGCTGTGACCGAGAGATCACTATGTTAGACAGAGATCTTAAGC 578  
Db 1646 CACAGCAGCAGAGAGCCACAGCAGCGGGAGCCACAGAGCGGGAGCCACAGCAGCGGGAGC 1705  
Qy 579 TGAAGAAGAGCAGCAGCTGTGAGAGAGGCTCCAAAGCCGCCAGCTTGAGAGC 638  
Db 1706 CACAGCAGCAGGAGCCACAGCAGCGGGAGCCACAGCAGCGGGAGCCACAGCAGCGGGAGC 1765  
Qy 639 CCGTGTCAACCGCGCCCATCTGAGTCCAGAGCAGCGCTGTGAGATCATCTAGAGC 698  
Db 1766 CACAGCAGCAGGAGCCACAGCAGCAGAGATGAGCAGCAGGATGAGCAGCAGCAGAGATG 1825  
Qy 699 AGAACCAGAAAGAGCTGAGCTGCACTCGAATCTTGAAAGCCTGAGGCCCCAGAGTG 758  
Db 1826 AGCAGCAGCAGGATGAGCAGCAGCAGAGATGAGCAGCAGGATGAGCAGCAGCAGAGATG 1885  
Qy 759 AGCTGCGCTGTACACAGCAGCCTCCGACACCCGAGATGTCACTGAGAAATCAAAATTA 818

Db 1886 AGCAGCAGCAGATGACGACGACGAGATGACGACGACGATGACGACGACGATG 1945  
QY 819 ACCAGCGATGCGGAGAAAGCTAATCTTGTACTTCAAGAGAGAAATCAAGCTCGGAAC 878  
Db 1946 AGCAGCAGCAGATGACGACGACGAGATGACGACGAGATGACGACGAGATG 2005  
QY 879 AATGGAAGGAAAGTTCTCCGACGCTATGACCAAGTTCATGGAAGCCTTTGAAAAAAG 938  
Db 2006 AGCAGCAGCAGATGACGACGACGAGATGACGACGACGAGATGACGACGAGATG 2065  
QY 939 TGGAGCGATCGAAAAACAACCCGCGCGCGGCGCAAGAGAGCAAGTGGCGGAGTACT 998  
Db 2066 AGCAGCAGCAGATGACGACGACGAGATGACGACGAGATGACGAGATGACGAGAC 2125  
QY 999 ACGAAAGCAGTTCCCTGAGATCCGCAAGCAGCGCAGCTGCAAGAGCCGATGCAAGCA 1058  
Db 2126 AGGATGAGCAG-----GAGCAGCAGAGATGACAGCAGAGATGACGACGACGAG 2179  
QY 1059 GGGTGGGCGACGGGGGCACTGGCTGTCCATGTCGGCCGCGCAGCCGACAGAGTGT 1118  
Db 2180 ATGACGACGACGACGACGAGATGACGACGACGACGACGATGACGACGAGATGAGC 2239  
QY 1119 CAGAGATCATGATGAGCTTCTGAGCAGAGAACTTGGAAAGCAGATGCGCAGCTGG 1178  
Db 2240 AGCAGCAGCAGATGACGACGACGACGACGAGAGAGAGAGAGAGAGAGAGAGAGC 2299  
QY 1179 CCGTATCCCGCCCATGCTGTGTAACAAGCTGACAGCAGCAGCCATCAAGTTCAATCA 1238  
Db 2300 AGGAGTTAGAGAGAGCAGAGAGAGTTAGAGATCAGAGAGAGAGTTAGAGAGAGCAG 2359  
QY 1239 ACGGCTTATGGCCCAACCCCATGAAAGGTGTACAAAGACCGCAGGTCATGAACATGTGA 1298  
Db 2360 AGCAGAGATTAGAGAGAGCAGAGAGAGATTAGAGAGAGAGAGAGAGAGAGAGAGC 2419  
QY 1299 GTGAGCAGAGAAAGGACCTTCCGGGAGAGATTCAATGAGATCCCAAGACTTTG--- 1355  
Db 2420 AGGAGCAGAGATTAGAGAGAGAGAGAGAGAGATTAGAGAGAGAGAGAGAGATTAGAG 2479  
QY 1356 ---GCCCTGATCATCTTCTGTGAGAGAGAGACAGTGGCTGATGTCCTCTTATTACT 1412  
Db 2480 AGCAGAGAGAGAGATTAGAGAGAGAGAGAGAGAGAGATTAGAGAGAGAGAGAGAGC 2539  
QY 1413 ACCTGATTAAGAAATGAGAACTTATAGAGCTGTGTGAGAGAGCTATCGGCGCGCG 1472  
Db 2540 AAGACAGAGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 2599  
QY 1473 GCAAGAGCAGAGCAACAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1532  
Db 2600 AGGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 2659  
QY 1533 CCATGCCCCCGCAGCAGCCAGAGAGAGAAAGATGAGAGAGAGAGAGAGAGAGAGAG 1592  
Db 2660 AGGAGAGAGAGAGAGTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 2719  
QY 1593 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1652  
Db 2720 AAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 2779  
QY 1653 A 1653  
Db 2780 A 2780

RESULT 10  
US-09-410-399-1  
; Sequence 1, Application US/09410399  
; Patent No. 6482587  
; GENERAL INFORMATION:  
; APPLICANT: Robertson, Erle S.  
; APPLICANT: Corbett, Murray A.  
; TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA  
; FILE REFERENCE: UM-03778

; CURRENT APPLICATION NUMBER: US/09/410,399  
; CURRENT FILING DATE: 1999-10-01  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3489  
; TYPE: DNA  
; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-410-399-1  
Query Match 1.3%; Score 107.8; DB 3; Length 3489;  
Best Local Similarity 43.8%; Pred. No. 1.3e-10;  
Matches 631; Conservative 0; Mismatches 792; Indels 18; Gaps 3;  
QY 219 AGCCCGGAGATGAACGATCCAGAGAGCTTCACTTCGCGCAGAGTCCCATCTATACCTGC 278  
Db 1352 AGCCACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1411  
QY 279 CCGAGCTGGGGAAGTCAAGATGAGATTCAATTGAAAGCAAGCGCCCTGGGCTTGAAGCTGC 338  
Db 1412 CACAGCAGCAGAGAGCAGACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1471  
QY 339 TGGCTGACCCCTGCTGGAGCCGTGACCCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGC 398  
Db 1472 AGCAGCAGAGAGCTCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1531  
QY 399 AAGACTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 458  
Db 1532 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1585  
QY 459 CCGCGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 518  
Db 1586 CACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1645  
QY 519 TCCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 578  
Db 1646 CACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1705  
QY 579 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 638  
Db 1706 CACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1765  
QY 639 CCGTGTACCCGCGCCATTCAGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 698  
Db 1766 CACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1825  
QY 699 AGAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 758  
Db 1826 AGCAGCAGCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1885  
QY 759 AGCTGCGCTGTACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 818  
Db 1886 AGCAGCAGCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1945  
QY 819 ACCAGCGATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 878  
Db 1946 AGCAGCAGCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 2005  
QY 879 AATGGAAGGAAAGTTCTCCGACGCTATGACCAAGTTCATGGAAGCCTTTGAAAAAAG 938  
Db 2006 AGCAGCAGCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 2065  
QY 939 TGGAGCGATCGAAAAACAACCCGCGCGCGGCGCAAGAGAGCAAGTGGCGGAGTACT 998  
Db 2066 AGCAGCAGCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 2125  
QY 999 ACGAAAGCAGTTCCCTGAGATCCGCAAGCAGCGCAGCTGCAAGAGCCGATGCAAGCA 1058  
Db 2126 AGGATGAGCAG-----GAGCAGCAGAGATGACAGCAGAGATGACGACGACGAGAG 2179  
QY 1059 GGGTGGGCGACGGGGGCACTGGCTGTCCATGTCGGCCGCGCAGCCGACAGAGAGAGAGTGT 1118  
Db 2180 ATGAGCAGCAGCAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGGC 2239

QY 1119 CAGAGATCATCGATGCGCTCTCAGACGAGAGAACTTGAGAAAGATGCGCCAGCTGG 1178  
DB 2240 AGCAGAGACAGATGAAACAGAGACGACAGAGAGAGAGAGAGAGAGAGAGAGAG 2299  
QY 1179 CCGTATCCCGCCCATGCTGTACAGACGTGACGACGACGACATCAAGTTTCATCAATGA 1238  
DB 2300 AGAGATTAGAGAGACGAGAGAGAGATTAGAGATCAGAGACGAGAGATTAGAGAGACAG 2359  
QY 1239 AGGGGCTTATGCGCCGACCCCATGAAAGTTTACAAACCCCGCATGTCATGAAATTTGGA 1298  
DB 2360 AGCAGAGATTAGAGAGACGAGAGAGAGATTAGAGAGAGAGAGAGAGATTAGAGAGAGC 2419  
QY 1299 GTGACGAGAGAGAGAGACCTTCCGAGAGAACTTCATGACATCCCAAGACTTTG--- 1355  
DB 2420 AGAGAGAGAGAGATTAGAGAGACGAGAGAGAGATTAGAGAGAGAGAGAGAGATTAGAG 2479  
QY 1356 ---GCCTGATCGCATCTTCTTGAGAGAGAAAGACAGTGGCTGATGCTCTTATTACT 1412  
DB 2480 AGCAGAGACGAGAGATTAGAGAGACGAGAGAGATTAGAGAGACGAGAGAGATTAGAGAG 2539  
QY 1413 ACTGACTAAGAAAGATGAACTATTAAGAGCTGTGAGACGAGACTATCGCGCCGCG 1472  
DB 2540 AAGAGACAGAGAGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 2599  
QY 1473 GCAAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1532  
DB 2600 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 2659  
QY 1533 CCATGCCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1592  
DB 2660 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 2719  
QY 1593 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1652  
DB 2720 AAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2779  
QY 1653 A 1653  
DB 2780 A 2780

RESULT 11  
US-09-894-273-1  
; Sequence 1, Application US/09894273  
; Patent No. 6756203  
; GENERAL INFORMATION:  
; APPLICANT: Kieff, Elliott D.  
; APPLICANT: Bailestas, Mary B.  
; APPLICANT: Kave, Kenneth M.  
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO  
; FILE REFERENCE: 16412-10001R  
; CURRENT APPLICATION NUMBER: US/09/894,273  
; PRIOR FILING DATE: 2001-06-28  
; PRIOR APPLICATION NUMBER: US 60/109,422  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3489  
; TYPE: DNA  
; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-894-273-1

Query Match 1.3%; Score 107.8; DB 3; Length 3489;  
Best Local Similarity 43.8%; Pred. No. 1.3e-10;  
Matches 631; Conservative 0; Mismatches 792; Indels 18; Gaps 3;

QY 219 AGCCGCGAATGAAAGGTTCCGAGAGCTCACTGCGCCAGAGTCCCACTATACCTGC 278  
DB 1352 AGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1411

QY 279 CCGAGCTGGAGAAAGTCAAGATGAGAGTTCAATTGAAAGCAAGCCCTCGCTAGAGCTGC 338  
DB 1412 CACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1471  
QY 339 TGCCTGACCCCTGCTGCTGCAAGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 398  
DB 1472 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1531  
QY 399 AAGACTTCAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 458  
DB 1532 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1585  
QY 459 CCGCGACACTGACCCCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 518  
DB 1586 CACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1645  
QY 519 TCCAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 578  
DB 1646 CACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1705  
QY 579 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 638  
DB 1706 CACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1765  
QY 639 CCGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 698  
DB 1766 CACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1825  
QY 699 AAGAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 758  
DB 1826 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1885  
QY 759 AGCTGCGCTGTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 818  
DB 1886 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1945  
QY 819 ACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 878  
DB 1946 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2005  
QY 879 AATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 938  
DB 2006 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2065  
QY 939 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 998  
DB 2066 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2125  
QY 999 AGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1058  
DB 2126 AGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2179  
QY 1059 GGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1118  
DB 2180 ATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2239  
QY 1119 CAGAGATCATGATGAGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1178  
DB 2240 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2299  
QY 1179 CCGTATCCCGCCCATGCTGTACAGACGTGACGACGACGACATCAAGTTTCATCAATGA 1238  
DB 2300 AGGAGATTAGAGAGACGAGAGAGAGATTAGAGATCAGAGACGAGAGATTAGAGAGACAG 2359  
QY 1239 AGGGGCTTATGCGCCGACCCCATGAAAGTTTACAAACCCCGCATGTCATGAAATTTGGA 1298  
DB 2360 AGCAGAGATTAGAGAGACGAGAGAGAGATTAGAGAGAGAGAGAGAGATTAGAGAGAGC 2419  
QY 1299 GTGACGAGAGAGAGAGACCTTCCGAGAGAACTTCATGACATCCCAAGACTTTG--- 1355  
DB 2420 AGAGAGAGAGAGATTAGAGAGACGAGAGAGAGATTAGAGAGAGAGAGAGAGATTAGAGAG 2479  
QY 1356 ---GCCTGATCGCATCTTCTTGAGAGAGAAAGACAGTGGCTGATGCTCTTATTACT 1412





Db 19577 AGGAGCAGAGGTTAGAGCAGAGCAGAGGAGTTAGAGGAGCAGAGGAGTTAGAGG 19518  
QY 1356 ---GCTGATCGATCATCTTCTTGAGAGAGAGACAGTGTGCTGAGTCTCTTATTA 1412  
Db 19517 AGCAGAGCAGAGGTTAGAGAGCAGAGCAGAGGAGTTAGAGAGCAGAGAGTGAAGAGC 19458  
QY 1413 ACCTGACTAAGAGAAATGAGAACTTAAAGACCTGTGTAGACCGAGCTATCGCGCCG 1472  
Db 19457 AGAGCAGAGAGCTGAGAGAGCAGAGCAGAGCAGAGAGAGGAAATTAGAGAGGTG 19398  
QY 1473 GCAAGAGCCAGCAGCAACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1532  
Db 19397 AGGAGCAGAGCAGAGCAGAGAGAGAGAGAGAGCAGAGGTTAGAGAGGTGGAAGAGC 19338  
QY 1533 CCATGCCCCCAGCAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1592  
Db 19337 AGGAGAGCAGAGGATTAGAGAGGTGAGAGAGCAGAGAGAGGTTAGAGAGGTG 19278  
QY 1593 AGGAGAGAGAGAGAGCCGAGAGGTGAGAGAGCAGAGAGAGCCTCTCAAGAGAGAGCAG 1652  
Db 19277 AGAGCAGAGAGCAGCAGAGGTTAGAGAGGTGAGAGAGCAGAGCAGAGAGGAGTGAAC 19218  
QY 1653 A 1653  
Db 19217 A 19217

RESULT 13  
US-08-757-669A-20/c  
Sequence 20, Application US/08757669A  
Patent No. 6183751  
GENERAL INFORMATION:  
APPLICANT: Chang, Yuan  
APPLICANT: Bohenzky, Roy A.  
APPLICANT: Russo, James J.  
APPLICANT: Edelman, Isidore S.  
APPLICANT: Moore, Patrick S.  
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS  
TITLE OF INVENTION: SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/757,669A  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 45185-F  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32207 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-757-669A-20

Query Match 1.3%; Score 107.8; DB 3; Length 32207;  
Best Local Similarity 43.8%; Pred. No. 2.5e-10;  
Matches 631; Conservative 0; Mismatches 792; Indels 18; Gaps 3;  
QY 219 AGCCGGGAATGACCGTCTCCAGAGACTCCACTTGCGCCAGAGTCCACTTACTTGC 278  
Db 20645 AGCCACAGCAGCGAGAGCCACAGCAGCAGAGAGCCCTGCGAGAGCCACACAGCAGAGC 20586  
QY 279 CCGAGCTGGGGGAAGTGAAGATGAGATTCAATTGAAAGCAAGCCCTCGGCTAGAGTGC 338  
Db 20585 CACAGCAGCAGAGCCACAGCAGCAGAGAGCCCTGCGAGAGCCACACAGCAGAGAGCCAC 20526  
QY 339 TGCCTGACCCCTGCTGCGACCGTCAACCCCTGTGCGCACCGGCAAGCCTGCGGATCTG 398  
Db 20525 AGCAGAGAGAGCCCTGCGAGAGCCACACAGAGAGCCACACAGAGAGCCACACAGAGCCACAGC 20466  
QY 399 AAGACTTACCAAGGACCGTAGCTTGAAGGCAAGTGTGAACCGGTGTCTCCCCCAAGCC 458  
Db 20465 AGCAGAGCCACAGCA-----GCAGAGCCACAGCAGAGAGAGCCACAGCAGCAGAGC 20412  
QY 459 CCCCCGACACTGACCTGAGCTGAGCTGTGCTGCGCCGACGCGCTGTCCAGAGAGAGCTGA 518  
Db 20411 CACAGCAGCAGAGCCACAGCAGCAGAGAGCCACAGCAGCAGAGCCACAGCAGCAGAGAGC 20352  
QY 519 TCCAGAACATGAGCCGCTGGAACCGAGAGATCACTATGTTAGAGCAGCAGAGATCTTAAGC 578  
Db 20351 CACAGAGCGGGAGCCACAGCAGCAGCGGAGCCCAAGAGCGGGAGCCACAGCAGCGGGAGC 20292  
QY 579 TGAAGAGAGAGCAGCAACAGCTGAGAGAGAGAGGCTCCAGAGCCGCGGAGCTGAAGAGC 638  
Db 20291 CACAGCAGCGGGAGCCACAGCAGCAGCGGGAGCCACAGCAGCGGGAGCCACAGCAGCGGAGC 20232  
QY 639 CCGTGTCAACCGCCGCCCATCGAGTGAAGCAGCAGCAGCTGTGTGAGATCATTAAGAGC 638  
Db 20231 CACAGAGCGGGAGCCACAGCAGCAGAGATGAGCAGAGATGAGCAGAGCAGCAGAGATG 20172  
QY 699 AGAACCGGAAGAGCTGAAGCTGACATCGATCGATTCTGGAAGGCTGAGGAGCCCGCAGGTG 758  
Db 20171 AGCAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGCAGATGAGCAGCAGCAGAGATG 20112  
QY 759 AGCTGCGCTGTACAAACAGCCCTCCGACACCCGCGCATATCATGAAACATCAAAATTA 818  
Db 20111 AGCAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGAGATG 20052  
QY 819 ACCAGGCAATGCGGAAGAGCTTAATTTGTACTTCAAGAGAGAGATCAACGCTCGGAAGC 878  
Db 20051 AGCAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAGCAGAGATG 19992  
QY 879 AATGGAAGCAGAAATTCTGCGCAGCGCTATGACCAAGCTCATGAGAGCCTTGGAAAAAAGG 938  
Db 19991 AGCAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAGCAGAGATG 19932  
QY 939 TGGAGCGCATCGAAAAACCCGCGCGCGGCGGCGCAAGAGAGCAAGGTGCGGAGTACT 998  
Db 19931 AGCAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGAGCAGCAGATGAGCAGAGCAGAG 19872  
QY 999 ACGAAAGCAGTTTCCCTGAGATCCGACAGCAGCGCAGCTGCGAGAGCCGATCAAGACA 1058  
Db 19871 AGGATGAGCAG-----GACAGCAGAGATGAGCAGCAGAGATGAGCAGCAGCAGCAG 19818  
QY 1059 GGGTGGGCGCAGCGGGCAGTGGGCTGTCCATGTGCGCCCGCCAGCAGCAGCAGAGGTG 1118  
Db 19817 ATGAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGCAGATGAGCAGCAGCAGAGATGAGC 19758  
QY 1119 CAGAGATCATGATGCGCTTCAAGCAGAGAACTTGAAGAAAGCAGATGCGCAGCTGG 1178  
Db 19757 AGCAGCAGCAGATGAAACAGAGCAGCAGAGAGAGAGCAGAGCAGCAGAGAGC 19698  
QY 1179 CCGTATCCCGCCCATGCTGTGAGAGCTGACAGCAGAGCAGCATCAAGTTTCAATCAATGA 1238  
Db 19697 AGGAGTTAGAGAGCAGAGCAGAGGAGTTAGAGATTCAGAGCAGAGGATTAAGAGAGCAG 19638  
QY 1239 ACGGCTATGAGCCGACCCCATGAAGGTGTACAAAGACCGCCAGGTCATGAACATGTGA 1298

Db 19637 AGCAGAGTTAAGAGACGAGGACGAGAGTTAGAGAGGACGAGAGTTAGAGAGC 19578  
QY 1299 GTGAGCAGAGAGAGACCTTCCGAGAGAGTTGATGACGATCCAGAACTTTG--- 1355  
Db 19577 AGAGCAGAGAGTTAGAGAGCAGAGCAGAGTTTAGAGAGCAGAGAGTTAGAG 19518  
QY 1356 ---GCCGTATCGCATCATCTTCTTGAGAGAGAGACAGTGGCTGATGTCCTCTATTACT 1412  
Db 19517 AGCAGAGCAGAGAGTTAGAGAGCAGAGAGTTAGAGAGCAGAGAGTTAGAGAGC 19458  
QY 1413 ACCTGACTTAAGAGAGATGAGACTATAGAGCTGTGAGAGCGAGCTATCGGCCCGG 1472  
Db 19457 AAGACAGAGAGGTGAGAGCAGAGCAGAGCAGAGAGCAGAGATTAGAGAGGTGG 19398  
QY 1473 GCAGAGCCAGAGCAACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1532  
Db 19397 AGAGCAGAGAGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 19338  
QY 1533 CCATGCCCCGACGACGACGAGAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAG 1592  
Db 19337 AGGAGAGCAGAGAGTTAGAGAGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 19278  
QY 1593 AGGAGAGAGAGAGAGCGGAGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1652  
Db 19277 AAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 19218  
QY 1653 A 1653  
Db 19217 A 19217

RESULT 14  
US-09-230-371A-20/c  
; Sequence 20, Application US/09230371A  
; Patent No. 6348586  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Yuan  
; APPLICANT: Bohenzky, Roy A  
; APPLICANT: Russo, James J  
; APPLICANT: Edelman, Reidore S  
; APPLICANT: Moore, Patrick S  
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: 45185-G-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/230,371A  
; PRIOR FILING DATE: 1999-11-17  
; PRIOR APPLICATION NUMBER: PCT/US97/13346  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 32207  
; TYPE: DNA  
; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-230-371A-20

Query Match 1.3%; Score 107.8; DB 3; Length 32207;  
Best Local Similarity 43.8%; Pred. No. 2.5e-10;  
Matches 631; Conservative 0; Mismatches 792; Indels 18; Gaps 3;

QY 219 AGCCCGGAGATGAACGCTCCAGAGCTCCAGCTGCGCCAGAGTCCCATCTATACCTGC 278  
Db 20645 AGCCACAGACAGAGAGCCACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20586  
QY 279 CCGAGCTGGGAGAGTCAAGATGAGATTATTTGAAGCAAGCGCCCTCGGCTGAGAGCTGC 338  
Db 20585 CACACACAGAGAGAGCCACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20526  
QY 339 TGCCTGACCCCTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 398  
Db 20525 AGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20466

QY 399 AAGACTTACCAAGAGACCGTGAAGCTTGAAGGCGCAAGCTGGAACCGGTGTCTCCCCCAGCC 458  
Db 20465 AGCAGAGAGCCACAGCA-----GAGAGAGCCACAGACAGAGAGAGAGAGAGAGAGAGC 20412  
QY 459 CCCCAGCAGTGAACCTGAGCTGAGAGCTGTGTCGCCACCGCTGTCCAAGAGAGAGTGA 518  
Db 20411 CACAGCAGAGAGAGCCACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20352  
QY 519 TCCAGAACATGAGACCGCTGAGAGCCGAGAGATCACATGATGAGAGAGAGATCTTAAAGC 578  
Db 20351 CACAGCAGCGGAGAGCCACAGCAGCGGAGAGCCACAGCAGCGGAGAGCCACAGCAGCGGAGC 20292  
QY 579 TGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 638  
Db 20291 CACAGCAGCGGAGAGCCACAGCAGCGGAGAGCCACAGCAGCGGAGAGCCACAGCAGCGGAGC 20232  
QY 639 CCGGTGCAACCGCGCCCATCGATCGATGAGAGCAACCGAGCTGTGAGATCATCTACAGC 698  
Db 20231 CACAGCAGCGGAGAGCCACAGCAGCGGAGAGCCACAGCAGCGGAGAGCCACAGCAGCGGAGC 20172  
QY 699 AAGAACCGAAGAGAGCTGAAGCTGCAATCGGATTTTGAAAGGCTTGAGGCCCAAGTGG 758  
Db 20171 AGCAGCAGCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20112  
QY 759 AGCTGCGCTGTACAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 818  
Db 20111 AGCAGCAGCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20052  
QY 819 ACCAGGCGATCGGAGAGAGCTAATCTTGTACTTCAAGAGAGAGAGATCAGCTCGGAAGC 878  
Db 20051 AGCAGCAGCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19992  
QY 879 AATGGAAGCAGAGATTTGCGAGCGCTATGACAGCTCATGAGAGCTTGGAAGAGAGAG 938  
Db 19991 AGCAGCAGCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19932  
QY 939 TGAAGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 998  
Db 19931 AGCAGCAGCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19872  
QY 999 AGGAAAGCAGATTCCTGAGATCCGACAGCAGCGGAGAGCTGAGAGAGAGAGAGAGAGAG 1058  
Db 19871 AGGATGAGAGAG-----GAGCAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 19818  
QY 1059 GGGTGGGCGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1118  
Db 19817 ATGAGCAGCAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19758  
QY 1119 CAGAGATCATGATGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1178  
Db 19757 AGCAGCAGCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19698  
QY 1179 CCGGATCCCGCCCATGCTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1238  
Db 19697 AGGAGTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19618  
QY 1239 AGGGCTTATGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1298  
Db 19637 AGCAGAGATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19578  
QY 1299 GTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1355  
Db 19577 AGGAGCAGAGAGTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19518  
QY 1356 ---GCCGTATCGCATCATCTTCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1412  
Db 19517 AGCAGAGCAGAGAGTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19458  
QY 1413 ACCTGACTTAAGAGAGATGAGACTATAGAGCTGTGAGAGCGAGCTATCGGCCCGG 1472  
Db 19457 AAGACAGAGAGGTGAGAGCAGAGCAGAGCAGAGCAGAGAGAGAGAGAGAGAGAGAGAG 19398  
QY 1473 GCAGAGCCAGAGCAACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1532  
Db 19397 AGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 19338  
QY 1533 CCATGCCCCGACGACGACGAGAGAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAG 1592  
Db 19337 AGGAGAGCAGAGAGTTAGAGAGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19278  
QY 1593 AGGAGAGAGAGAGAGCGGAGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1652  
Db 19277 AAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 19218  
QY 1653 A 1653  
Db 19217 A 19217



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2006, 05:34:50 ; Search time 27353 Seconds  
(without alignments)  
17811.765 Million cell updates/sec

Title: US-09-522-753-4

Perfect score: 8564.8  
Sequence: 1 catctgcggcctccacacagc.....caaaaaaaaaaaaaaaaaaa 8571

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Geneml:\*

1: gb\_ba:\*

2: gb\_in:\*

3: gb\_env:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pac:\*

7: gb\_ph:\*

8: gb\_dr:\*

9: gb\_ro:\*

10: gb\_sbs:\*

11: gb\_sy:\*

12: gb\_un:\*

13: gb\_vl:\*

14: gb\_hcg:\*

15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8541	99.7	8561	8 AF113003	AF113003 Homo sapi
2	8324.6	97.2	8666	8 AF125672	AF125672 Homo sapi
3	8226.8	96.1	9053	6 AR447713	AR447713 Sequence
4	8038.6	93.9	8548	8 AY965853	AY965853 Homo sapi
5	5455	63.7	5989	8 HSU37146	U37146 Human sllen
6	5037.2	58.8	8544	9 AF125671	AF125671 Mus muscu
7	4981.6	58.2	8388	9 AF113001	AF113001 Mus muscu
8	4378.2	51.1	7465	9 AF113002	AF113002 Mus muscu
9	4253.8	49.7	4686	8 AB209089	AB209089 Homo sapi
10	3963.8	46.3	6339	6 CQ722208	CQ722208 Sequence
11	2759.8	32.2	2842	8 BC004326	BC004326 Homo sapi
12	2618.6	30.6	2930	8 BC047524	BC047524 Mus muscu
13	1703.4	19.9	2964	9 BC073916	BC073916 Homo sapi
14	1132.4	13.2	205283	8 AC027706	AC027706 Homo sapi
15	948.4	11.1	161970	14 AC027706	AC027706 Homo sapi
16	898.4	10.5	956	6 BD270289	BD270289 Human nuc
17	898.4	10.5	956	6 AX035226	AX035226 Sequence
18	898.4	10.5	956	6 AX342080	AX342080 Sequence

19	898.4	10.5	956	6 AX342140	AX342140 Sequence
20	898.4	10.5	956	6 AX441143	AX441143 Sequence
21	856.4	10.0	1907	5 CR926317	CR926317 Xenopus t
22	854.6	10.0	1896	5 BC082706	BC082706 Xenopus t
23	854.6	10.0	1917	5 BC054296	BC054296 Xenopus 1
24	732.8	8.6	79494	14 AC068837	AC068837 Homo sapi
25	652.8	7.6	752	6 CQ769363	CQ769363 Sequence
26	650	7.6	650	6 AX677743	AX677743 Sequence
27	631.6	7.4	7949	8 AB028970	AB028970 Homo sapi
28	630	7.4	7940	6 CQ847916	CQ847916 Sequence
29	630	7.4	7940	6 AX578069	AX578069 Sequence
30	630	7.4	7940	6 AF044209	AF044209 Homo sapi
31	629.6	7.4	780	9 MMU5312	U35312 Homo sapi
32	626.8	7.3	3120	8 AF303586	AF303586 Homo sapi
33	626.8	7.3	8018	8 AF087856	AF087856 Homo sapi
34	615.8	7.2	7264	9 BC086657	BC086657 Mus muscu
35	572.2	6.7	8626	5 BC099620	BC099620 Xenopus t
36	570	6.7	725	10 BV448749	BV448749 NCOR2_795
37	568.4	6.6	8959	5 AF495886	AF495886 Xenopus 1
38	555	6.5	555	6 AX677866	AX677866 Sequence
39	554.8	6.5	560	8 HSU80761	U80761 Homo sapien
40	553.6	6.5	560	8 HSU80750	U80750 Homo sapien
41	542.4	6.3	718	6 AX753058	AX753058 Sequence
42	540.6	6.3	1891	8 BC050594	BC050594 Homo sapi
43	534	6.2	534	6 AX778323	AX778323 Sequence
44	533.4	6.2	1850	8 BC058511	BC058511 Homo sapi
45	533.2	6.2	1808	8 BC068996	BC068996 Homo sapi

## ALIGNMENTS

RESULT 1	AF113003	8561 bp	mRNA	linear	PRI 20-MAR-1999
LOCUS	AF113003				
DEFINITION	Homo sapiens silencing mediator of retinoic acid and thyroid hormone receptor alpha mRNA, complete cds.				
ACCESSION	AF113003				
VERSION	AF113003.1	GI:4454551			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 8561)				
AUTHORS	Ordentlich,P., Downes,M., Xie,W., Genin,A., Spinner,N.B. and Evans,R.M.				
TITLE	Unique forms of human and mouse nuclear receptor corepressor SMRT				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (6), 2639-2644 (1999)				
PubMed	10077563				
REFERENCE	2 (bases 1 to 8561)				
AUTHORS	Downes,M.R., Ordentlich,P. and Evans,R.M.				
TITLE	Direct Submission				
JOURNAL	Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA				
FEATURES	Location/Qualifiers				
source	1..8561				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/chromosome="12"				
	/map="12q24"				
	/cuiuse_type="pituitary"				
	2..7555				
	/function="transcriptional co-repressor"				
	/note="hSMRT alpha; longer isoform than previously reported"				
	/codon_start=1				
	/product="silencing mediator of retinoic acid and thyroid hormone receptor alpha"				
	/protein_id="AAD20946.1"				



/db\_xref="GI:4454552"  
/translation="MSGTQVLAQWTRATERYPPHSLSYPIQIARTHTDVLGYOH  
HSRDYASHLSIPSI1IQORRRPSLISBPQNGERSQELHRESHSYLELKSMEIF  
IESKRRLLELPDLRPLATGQAPASIEDLTDORSJLTGLBEPSPSPHTPBL  
ELVPRLSKEELIOMNDYRDEITWVQOISLTKKKOQDLEBAKAPPEKVPBP  
IESKRSLSVOI1YDENRKAENARH1LEBGLQOVELPLRNSDPROYHENTK1QOM  
BKLIYPRRPHARKKOKOKFCORYDOLMBALBEPKVR1ENPNPRAYESVRYOY  
KOPPEIRKORELOERMOSHVGORGSLSNARSEHVESEITDGLSEONENKQKOL  
AVIPMLYDADQOR1KFINMGMLMADPMKLVYDRQVMNMSQEKSETEFEKMOHPN  
FOLIASFLERKTVACVLVYLTKNKENTKSLVRSYRARGSQOQOQOQOQOQOQ  
QOPMRSQSEDEKKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKE  
KGRKTANSQGRKGR1ITREMANEANSBEA1ITPOQSAELASMLNSENRETEEMAK  
KGLLEHGRNWSA1IARNVSKTYSCKN1FVFNK3KONDEL1IQOKH1KHEKRNARK  
KKGAFAASBEAPPPVEDEMEASGVSGNEWESEBEALHASGNEVPRSECPA  
TVNNSDYES1ESPHTEAKDQNGQPKPATLGADGPPGPTPRRTSRAP1EPT  
KAEATGAPTPPAPSPAPPPVPAVEKEBEETAAP1PVESEBOKPRAEELADVT  
LSEBVEKSECTEEAEEGPAKGDVAEAEAGALKAKEKEGSGAGTAKESGAPOD  
SSSASTGADVEDBAEGDKNRLSPRBL1ITGDPBRNAPASPOPLDKOLKORAAA  
1PR1QVTKYHREPREDAATKRAPRAPPRPQMLQPSBAPQOQSPKRSKPAPPA  
DKEATPAEAKQLEDPKPCWTSGLPFPFPREVITKASPHADPSATYAPPGHPLGL  
HDTAPVLYRPTL1SNRPLISSAKHPSVLERQIGAI1SQGKSVOLHVPSEHAKAVG  
PVTMG1PLPMDKRLAPFSGVQEQOLSPPRQAPPSLSLQVPAQESVLRGTALGAVP  
GSGITKGI1PSTVPDSAITVHGS1THGPADVLYKGTTR1IGBDSPELDRGSDS  
LPGHVTYEGKKGHVLSYEGKMSVTCSEKEDRSSGPHETAA1PKRTYDMEGVGR  
AISAS1SIGLMGRAL1PREHSPHNLKEOH1TIGS1ITQCI1PBYVAQEDVY1ABAKL  
KREGTPPPPPSRDL1TEAKTLOGLK1KRAHEGLVATVKAAGS1HE1PBE1RHT  
PEL1APRPLKES1ITQETPLKYDTGASTGSKHDVRL1GSPRTFPEVPLR1LMA  
DARALERCYBESLKSPPGTASSSGS1ARGAPV1PVE1GKRGOS1LYEDGASVP  
HLRGSVP1MPTPRLOGLSSSKASODRLKLTSTPRE1ASPH1VTEHHPH1SP  
YEH1LRGVSGVDLYRSH1PLADPTS1PBG1PLDAAA1YL1RHL1A1NPTYH1L1PPY  
L1RGP1DPA1LENROTT1INDY1TTSOOMHN1TMA1QARAD1RGLSPRESS1ALNYAA  
GREG1IDSQVPH1VLPPTGTPYAMDRLAV1PLAP1QDPSHRS1SS1LPGATTH  
LTKPTT1SSSERERDRERDRERREKSL1N1STTVBERAP1WRGETBSSSSSSG  
GGGSSSRPASHSHOHSP1SPRTQDALQORPSVLHMGMI1ITAVEPSP1TVLRS  
T1SGVPAP1ATP1PATK1PLGGLDGV1PTMEPVL1P1KEMAR1VAPRAP1DQV  
FLAKPARSGLEPASPSPKSGSPR1PLVPSGHAT1TAR1PRA1A1PHAS1PPAP1PA  
SASDHRK1EKTOSKPS1OELER1SLGHNSS1SPGVRPSVSS1SLTHDGLK1KHL  
BELD1SH1EBEL1P1KOP1FVK1GGE1NAH1L1PL1R1P1ES1OP1SS1PL1QTA1P1GK1QORV  
V1LAOH1SEV1ITQDYTRHHP1QOL1SAP1L1PAP1LS1FPA1SG1P1AD1L1R1P1PD1L1P1PDH  
GAPARGSP1SGE1GKRS1PEPNKTS1VLG1GG1ED1G1E1PVS1P1EGMTE1PGHS1R1AVP1PL1YR1D  
G1OTE1P1R1MKS1PGNT1SOP1PAP1F1SK1L1TBS1N1AMV1SK1COE1INK1K1L1N1R1RE1P1ND  
SOP1GTE1FM1P1A1ITG1G1MTV1R1SOA1V1OEH1A1STM1G1E1A1I1R1A1L1M1K1G1YOM1E1SP1S  
ANA1NP1N1AS1A1SL1P1A1MP1T1A1D1G1SDH1L1T1SP1G1G1AK1R1A1G1R1SS1R1K1A1S1P1AG1A  
SD1RP1SV1SV1S1B1D1CN1R1P1L1N1R1W1ED1R1PS1AG1ST1P1F1N1P1L1M1R1Q1A1V1M1AS1P1P  
PGL1P1AG1S1D1P1A1G1PH1H1AM1DE1BK1PL1CS1QY1ET1L1D1SE"

ORIGIN

Query Match	99.7%	Score 8541;	DB 8;	Length 8561;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 8560;	Conservative 1;	Mismatches 0;	Indels 10;	Gaps 1;
QY	1	CATGTCGAGCTCCACACACAGCTTGTGGCAGACAGTGGAGGAGCCACTGAGCCCGCTACCC	60	
DB	1	CATGTCGAGCTCCACACAGCTTGTGGCAGACAGTGGAGGAGCCACTGAGCCCGCTACCC	60	
QY	61	GCCCCACAGCCTTTTCTCTACCCAGTGCAGATCGCCCGGACGCAACGGAAGTGGGCTCTT	120	
DB	61	GCCCCACAGCCTTTTCTCTACCCAGTGCAGATCGCCCGGACGCAACGGAAGTGGGCTCTT	120	
QY	121	GGAAGTACAGACCACTCCGCGAGCTATGCTCCCACTGTGCGCGGGCTTCATATCA	180	
DB	121	GGAAGTACAGACCACTCCGCGAGCTATGCTCCCACTGTGCGCGGGCTTCATATCA	180	
QY	181	GCCCAAGCGGCGGAGGCTCTCCCTGCTGTGTGAGTTCCAGCCCGGGAATGAACGTCCTCA	240	
DB	181	GCCCAAGCGGCGGAGGCTCTCCCTGCTGTGTGAGTTCCAGCCCGGGAATGAACGTCCTCA	240	
QY	241	GGAAGTCCACCTGCGCGCAGAGTCCCATCTCATACCTGCGGAGCTGGGAAAGTCAAGAT	300	
DB	241	GGAAGTCCACCTGCGCGCAGAGTCCCATCTCATACCTGCGGAGCTGGGAAAGTCAAGAT	300	
QY	301	GGAAGTTCAATTGAAAGGAAGCGCCCTCGGCTAAGAGCTGCTGACCCCTCTGCTGACAC	360	
DB	301	GGAAGTTCAATTGAAAGGAAGCGCCCTCGGCTAAGAGCTGCTGACCCCTCTGCTGACAC	360	

DB	301	GGAAGTTCAATTGAAAGGAAGCGCCCTCGGCTAAGAGCTGCTGACCCCTCTGCTGACAC	360	
QY	361	GTACCCCTCTGTGCTGCACAGGCGCAAGCTTGGCGGATCTGAAGACTTCCACGAAGACCTTAG	420	
DB	361	GTACCCCTCTGTGCTGCACAGGCGCAAGCTTGGCGGATCTGAAGACTTCCACGAAGACCTTAG	420	
QY	421	CTTGAACGGGCAAGCTGGAAACCGGTGTCCTCCCGCAGCCCGCGCACTGACACCTGAGGT	480	
DB	421	CTTGAACGGGCAAGCTGGAAACCGGTGTCCTCCCGCAGCCCGCGCACTGACACCTGAGGT	480	
QY	481	GGAAGTGTGCGCGCACGCGCTGTCCAAAGAGAGAGTGTATCCAGAACATGACACCGGTGGA	540	
DB	481	GGAAGTGTGCGCGCACGCGCTGTCCAAAGAGAGAGTGTATCCAGAACATGACACCGGTGGA	540	
QY	541	CCGAGAGATCAACATGTGTAGACAGACAGATCTTAAGCTGAAGAAAGACAGCAACAGCT	600	
DB	541	CCGAGAGATCAACATGTGTAGACAGACAGATCTTAAGCTGAAGAAAGACAGCAACAGCT	600	
QY	601	GGAAGAGAGAGCTGCGCAAGCCCGGACCTGAGAAAGCCCGTGTACCGCGCGCCATGGA	660	
DB	601	GGAAGAGAGAGCTGCGCAAGCCCGGACCTGAGAAAGCCCGTGTACCGCGCGCCATGGA	660	
QY	661	GTGGAAGCACCGCAAGCTGTGTCAGATCATCAAGACGAAACCGGAAGAAAGCTGAAGC	720	
DB	661	GTGGAAGCACCGCAAGCTGTGTCAGATCATCAAGACGAAACCGGAAGAAAGCTGAAGC	720	
QY	721	TGCACATCGGAATTGTGAAAGGCTTGGAGGCCAGGTGAGCTGCCGTGTACAACCAAGCC	780	
DB	721	TGCACATCGGAATTGTGAAAGGCTTGGAGGCCAGGTGAGCTGCCGTGTACAACCAAGCC	780	
QY	781	CTCCGACACCCCGGCAAGTATCATGAGAAATCAAAATTAACCGAGGAGTCCGGAAGAGCT	840	
DB	781	CTCCGACACCCCGGCAAGTATCATGAGAAATCAAAATTAACCGAGGAGTCCGGAAGAGCT	840	
QY	841	AATCTTGAATCTCAAGAGAGAGAAATCAAGCTGGAAGAAATGGAAGCAAGATTCTGCA	900	
DB	841	AATCTTGAATCTCAAGAGAGAGAAATCAAGCTGGAAGAAATGGAAGCAAGATTCTGCA	900	
QY	901	GCGCTATGACCAAGCTCATGAGAGGCTTTGGAAAAAAGGTGAGCGCATGAAAAACAACC	960	
DB	901	GCGCTATGACCAAGCTCATGAGAGGCTTTGGAAAAAAGGTGAGCGCATGAAAAACAACC	960	
QY	961	GCGCGCGCGGCGCAAGAGACGAAGGTGCGCGAGTACTACGAAAGCAAGTTCCGTGAAGT	1020	
DB	961	GCGCGCGCGGCGCAAGAGACGAAGGTGCGCGAGTACTACGAAAGCAAGTTCCGTGAAGT	1020	
QY	1021	CCGCAAGACGCGAGCTGTGACAGAGCGCATGCGAGAGCGAGGTGTCAAGATCATGATGGCTTTC	1080	
DB	1021	CCGCAAGACGCGAGCTGTGACAGAGCGCATGCGAGAGCGAGGTGTCAAGATCATGATGGCTTTC	1080	
QY	1081	GCTGTCCATGTTGAGCCGCGCCGAGCGAGGACGAGAGTGTCAAGATCATGATGGCTTTC	1140	
DB	1081	GCTGTCCATGTTGAGCCGCGCCGAGCGAGGACGAGAGTGTCAAGATCATGATGGCTTTC	1140	
QY	1141	AGAGCAGAGAAACCTTGAGAAAGCAATGCGCTGCGCGTGAATCCGCGCCATGCTGTGA	1200	
DB	1141	AGAGCAGAGAAACCTTGAGAAAGCAATGCGCTGCGCGTGAATCCGCGCCATGCTGTGA	1200	
QY	1201	GGAAGCTGACAGCAGCGAGCATCAAGTTTATCAACATGAAACGGGCTTATGAGCCGACAT	1260	
DB	1201	GGAAGCTGACAGCAGCGAGCATCAAGTTTATCAACATGAAACGGGCTTATGAGCCGACAT	1260	
QY	1261	GAAAGTGTACAAAGACCGCAGAGTCATGAACATGTGAGTGAAGCAGAGAAAGAGACCTT	1320	
DB	1261	GAAAGTGTACAAAGACCGCAGAGTCATGAACATGTGAGTGAAGCAGAGAAAGAGACCTT	1320	
QY	1321	CCGGAAGAAATTCAATGACAGCATCCCAAGAACTTTGGCTGATCGATCTTCTGGAAG	1380	
DB	1321	CCGGAAGAAATTCAATGACAGCATCCCAAGAACTTTGGCTGATCGATCTTCTGGAAG	1380	
QY	1381	GAAAGCAGTGTGAGTGTGCTCTATTATCACTGAACCTGAAGGAATGAGAACTATTA	1440	
DB	1381	GAAAGCAGTGTGAGTGTGCTCTATTATCACTGAACCTGAAGGAATGAGAACTATTA	1440	

QY	1441	GAGCCTG	GTGAG	CGAGACT	TTTCGG	CGCCGCG	CAGAGCC	CAGAC	CAAC	CAC	CAG	CAG	C	1500
Dp	1441	GAGCCTG	GTGAG	CGAGACT	TTTCGG	CGCCGCG	CAGAGCC	CAGAC	CAAC	CAC	CAG	CAG	C	1500
QY	1501	GCAGCAG	AGAG	CGACG	CGACG	CGACG	CGACG	CGACG	CGACG	CGACG	CGACG	CGACG	CGACG	1560
Dp	1501	GCAGCAG	AGAG	CGACG	CGACG	CGACG	CGACG	CGACG	CGACG	CGACG	CGACG	CGACG	CGACG	1560
QY	1561	AGATGAG	AGAG	GAGAA	AGAG	CGCGAG	GAGAG	GAGAG	GAGAG	CGCGAG	GTGAG	AGAA	1620	
Dp	1561	AGATGAG	AGAG	GAGAA	AGAG	CGCGAG	GAGAG	GAGAG	GAGAG	CGCGAG	GTGAG	AGAA	1620	
QY	1621	CGACAG	AGAG	AGCTCT	CTCAG	AGAG	AGAC	CGAC	CTCAG	AGAG	AGAC	CGAC	1680	
Dp	1621	CGACAG	AGAG	AGCTCT	CTCAG	AGAG	AGAC	CGAC	CTCAG	AGAG	AGAC	CGAC	1680	
QY	1681	GAAAGAG	GTGTG	GTCTT	CCAA	AGGCTCG	CAAA	CTCG	CAAC	CGAG	AGAC	CGCAAA	1740	
Dp	1681	GAAAGAG	GTGTG	GTCTT	CCAA	AGGCTCG	CAAA	CTCG	CAAC	CGAG	AGAC	CGCAAA	1740	
QY	1741	CCGATCA	CCCGCT	CAAT	GGCTAA	TGAGG	CCAA	CAGCG	AGAG	GCAT	CA	CCCCCG	1800	
Dp	1741	CCGATCA	CCCGCT	CAAT	GGCTAA	TGAGG	CCAA	CAGCG	AGAG	GCAT	CA	CCCCCG	1800	
QY	1801	GAGCGCG	AGCTG	AGCTT	CCAT	AGAG	CTGAA	TGAG	AGTT	CTCG	CTT	GACAG	1860	
Dp	1801	GAGCGCG	AGCTG	AGCTT	CCAT	AGAG	CTGAA	TGAG	AGTT	CTCG	CTT	GACAG	1860	
QY	1861	GGAAA	CAGCC	CAAGAA	GGTCTT	CTT	GGAA	CA	CGG	CGG	CA	CTG	1920	
Dp	1861	GGAAA	CAGCC	CAAGAA	GGTCTT	CTT	GGAA	CA	CGG	CGG	CA	CTG	1920	
QY	1921	GGTGGG	CTCC	CAAG	CTGTG	TGCG	AGTGTAA	GA	ACTT	CTT	CA	CTT	1980	
Dp	1921	GGTGGG	CTCC	CAAG	CTGTG	TGCG	AGTGTAA	GA	ACTT	CTT	CA	CTT	1980	
QY	1981	GAACTCG	ATAGA	TCTTG	CGAC	AGCA	CAG	ACTGA	ATAGA	TGA	AGAA	AGAG	2040	
Dp	1981	GAACTCG	ATAGA	TCTTG	CGAC	AGCA	CAG	ACTGA	ATAGA	TGA	AGAA	AGAG	2040	
QY	2041	GAGGAG	AGAG	AGAA	GGCGG	CGGCGG	CGAG	CGAG	AGG	CTG	CA	TT	2100	
Dp	2041	GAGGAG	AGAG	AGAA	GGCGG	CGGCGG	CGAG	CGAG	AGG	CTG	CA	TT	2100	
QY	2101	GGAGAG	TGAG	AGAG	CGTGGG	GTGAG	CGTGA	ATAGA	TGA	AGAG	AGAG	AGAG	2160	
Dp	2101	GGAGAG	TGAG	AGAG	CGTGGG	GTGAG	CGTGA	ATAGA	TGA	AGAG	AGAG	AGAG	2160	
QY	2161	GGCTGAG	AGCTT	ATGCT	CTT	GGAG	ATAGA	GTGCC	CA	GAG	GGAG	ATG	2220	
Dp	2161	GGCTGAG	AGCTT	ATGCT	CTT	GGAG	ATAGA	GTGCC	CA	GAG	GGAG	ATG	2220	
QY	2221	CACCTG	CAAC	CAAGCT	CA	CA	CGAG	AGAT	CCCT	CT	CT	CA	2280	
Dp	2221	CACCTG	CAAC	CAAGCT	CA	CA	CGAG	AGAT	CCCT	CT	CT	CA	2280	
QY	2281	GGAC	CA	CAGG	CGA	ATGGG	CC	CA	AG	CC	CC	CA	2340	
Dp	2281	GGAC	CA	CAGG	CGA	ATGGG	CC	CA	AG	CC	CC	CA	2340	
QY	2341	AGGCGC	AC	CC	CA	CC	CA	CG	AG	AG	CA	CT	2400	
Dp	2341	AGGCGC	AC	CC	CA	CC	CA	CG	AG	AG	CA	CT	2400	
QY	2401	TGAAG	CA	CC	CG	AG	AG	AG	AG	AG	AG	AG	2460	
Dp	2401	TGAAG	CA	CC	CG	AG	AG	AG	AG	AG	AG	AG	2460	
QY	2461	TGTGTG	TC	CC	CA	AG	AG	AG	AG	AG	AG	AG	2520	
Dp	2461	TGTGTG	TC	CC	CA	AG	AG	AG	AG	AG	AG	AG	2520	

QY	2521	GGAGAGCAGAAAGCCCCCGGCTTGAGAGCTGGCATGTGACACAGAGAAAGCCGAGGA	2580
Db	2521	GGAGAGCAGAAAGCCCCCGGCTTGAGAGCTGGCATGTGACACAGAGAAAGCCGAGGA	2580
QY	2581	GCCGCTCAAGAGCGAGTGCACGAGAGGAAGCCGAGAGAGGGGCGCGCCAAAGGSCAAGAGCGC	2640
Db	2581	GCCGCTCAAGAGCGAGTGCACGAGAGGAAGCCGAGAGAGGGGCGCGCCAAAGGSCAAGAGCGC	2640
QY	2641	GGAAGCGCGTTGAGGCGCACGCGCCGAGGGGGCGCTCAAGCAGAGAGAGAGAGAGGCGGGAG	2700
Db	2641	GGAAGCGCGTTGAGGCGCACGCGCCGAGGGGGCGCTCAAGCAGAGAGAGAGAGAGGCGGGAG	2700
QY	2701	CGGCAAGGCGCCACCTGCGCAAGAGCTGGGGCGCCCCCAGGACAGCGCATCTCAATGTCTAC	2760
Db	2701	CGGCAAGGCGCCACCTGCGCAAGAGCTGGGGCGCCCCCAGGACAGCGCATCTCAATGTCTAC	2760
QY	2761	CTGCAAGTGCAGACGAGGTGTGATGAGGGCGGAGGGGGGGGACAAAGAACGGGCTGCTGTCCC	2820
Db	2761	CTGCAAGTGCAGACGAGGTGTGATGAGGGCGGAGGGGGGGGACAAAGAACGGGCTGCTGTCCC	2820
QY	2821	AAGGCCAGAGCCTCTTCAACCCCGACCTGGCGACCCCGGGGCGCAATGCTCAACCCAGAAAGCC	2880
Db	2821	AAGGCCAGAGCCTCTTCAACCCCGACCTGGCGACCCCGGGGCGCAATGCTCAACCCAGAAAGCC	2880
QY	2881	ACTGGACCTTGAAGCAGTGAAGCAGCAGCGGCTTCGATCCGCCCCCATTCAGAGTCAACCA	2940
Db	2881	ACTGGACCTTGAAGCAGTGAAGCAGCAGCGGCTTCGATCCGCCCCCATTCAGAGTCAACCA	2940
QY	2941	AGTTCATGAGCCCCCGGGGAGGACGCAAGCTTCCACCAAGCAGACTCCCCAGGCCAAC	3000
Db	2941	AGTTCATGAGCCCCCGGGGAGGACGCAAGCTTCCACCAAGCAGACTCCCCAGGCCAAC	3000
QY	3001	GCCACCGCAAAACCTTGACGCGGAGAGCGACGCCCCCAAGCAGCTTGGCAGACAGCAGCAGC	3060
Db	3001	GCCACCGCAAAACCTTGACGCGGAGAGCGACGCCCCCAAGCAGCTTGGCAGACAGCAGCAGC	3060
QY	3061	GGGCAAGAGCAGAGAGCCCGGACCCCGCGCGGACAAAGAGGGCTTTCGACGCCAGGCCCA	3120
Db	3061	GGGCAAGAGCAGAGAGCCCGGACCCCGCGCGGACAAAGAGGGCTTTCGACGCCAGGCCCA	3120
QY	3121	GAAAGCTGCTTGGGGACCCCCCTTGTGTGACATTCCGGGCTGCTTCCCCCGTGGCCCCCG	3180
Db	3121	GAAAGCTGCTTGGGGACCCCCCTTGTGTGACATTCCGGGCTGCTTCCCCCGTGGCCCCCG	3180
QY	3181	TGAGGTATCAAGGCTTCCCGGACATGCGCGGAGCCCTCAGCCTTCTCTCAAGCTTCAAC	3240
Db	3181	TGAGGTATCAAGGCTTCCCGGACATGCGCGGAGCCCTCAGCCTTCTCTCAAGCTTCAAC	3240
QY	3241	TGCTCACCCATGCGCCCTTGGGACCTTCATGACATGCGCCGCGCTCTTGCAGCGCCAC	3300
Db	3241	TGCTCACCCATGCGCCCTTGGGACCTTCATGACATGCGCCGCGCTCTTGCAGCGCCAC	3300
QY	3301	CACCATCTCCAACCCGCGCTCCCTCACTCTCTTGCACCAACACCCAGCGCTCTCGAGAG	3360
Db	3301	CACCATCTCCAACCCGCGCTCCCTCACTCTCTTGCACCAACACCCAGCGCTCTCGAGAG	3360
QY	3361	GCAAAATGAGTGCATCTCCCAAGGAATGTGATGATCAGACTCAAGTCCCGATCTAGAGCA	3420
Db	3361	GCAAAATGAGTGCATCTCCCAAGGAATGTGATGATCAGACTCAAGTCCCGATCTAGAGCA	3420
QY	3421	TGCGCAAGGCGCGGTGGGCTCTGTCAACATGAGGCTGCTCCTGCGCATGTGACCCCAAAA	3480
Db	3421	TGCGCAAGGCGCGGTGGGCTCTGTCAACATGAGGCTGCTCCTGCGCATGTGACCCCAAAA	3480
QY	3481	GCTTGCACCCCTTCAAGCGAGTGAAGCAGAGCAAGCTGTCCCAAGGGGCGCAGGCTTGGG	3540
Db	3481	GCTTGCACCCCTTCAAGCGAGTGAAGCAGAGCAAGCTGTCCCAAGGGGCGCAGGCTTGGG	3540
QY	3541	ACCGGAAAGCGCTGGGGGTGCGCCACAGCCACAGAGGCGCTGCTGTGAGAGGAAAGCTCT	3600
Db	3541	ACCGGAAAGCGCTGGGGGTGCGCCACAGCCACAGAGGCGCTGCTGTGAGAGGAAAGCTCT	3600
QY	3601	GGGCTCAGTTCGGGGCGAAGCATCACAAAGCATTCCAGACACAGGGTGCCTTGGAG	3660

Db 3601 GGGCTCAGTTCGGGCGGGAAGATCACTAAAGCATTCCACAGCACGCGTCCCTCCGA 3660  
QY 3661 CAGCGCATCATATACCGGCGCTCCATCACTCCAGCGGACCGAGCTGACGTCTGTACAA 3720  
Db 3661 CAGCGCATCATATACCGGCGCTCCATCACTCCAGCGGACCGAGCTGACGTCTGTACAA 3720  
QY 3721 GGGCACCATCACAGAGTATCGGCGAGGACAGCCGAGTCCGTTGACCCGCGCGGGA 3780  
Db 3721 GGGCACCATCACAGAGTATCGGCGAGGACAGCCGAGTCCGTTGACCCGCGCGGGA 3780  
QY 3781 GGAACAGCTGCCCAAGGGGCAAGTCTACGAAAGGCAAGAGGCGACAGTCTTGTCTTA 3840  
Db 3781 GGAACAGCTGCCCAAGGGGCAAGTCTACGAAAGGCAAGAGGCGACAGTCTTGTCTTA 3840  
QY 3841 TGAAGGTGATGTCTGTATCCAGTGTCCAAAGGAGACGGCAGAAAGGCTCAGAAC 3900  
Db 3841 TGAAGGTGATGTCTGTATCCAGTGTCCAAAGGAGACGGCAGAAAGGCTCAGAAC 3900  
QY 3901 CCCCATGAGACGGGCGCCCGCAAGCGCACCTATGACATATGAGAGGGCGCGTGGGAG 3960  
Db 3901 CCCCATGAGACGGGCGCCCGCAAGCGCACCTATGACATATGAGAGGGCGCGTGGGAG 3960  
QY 3961 AGCCATCTCTCAAGCCAGATGAAAGTCTCATGGGCGGTGCCATTCGCGCGGAGACA 4020  
Db 3961 AGCCATCTCTCAAGCCAGATGAAAGTCTCATGGGCGGTGCCATTCGCGCGGAGACA 4020  
QY 4021 CAGCCCCCAGCATCAAAAGAGAGACCAATCCGCGGGTCCATTCACAAAGGATCCC 4080  
Db 4021 CAGCCCCCAGCATCAAAAGAGAGACCAATCCGCGGGTCCATTCACAAAGGATCCC 4080  
QY 4081 TCGGTCTTACGTGAGGACACAGAGAGCTACCTGCGTCGGAGGCGCAAGCTCTTAAAGG 4140  
Db 4081 TCGGTCTTACGTGAGGACACAGAGAGCTACCTGCGTCGGAGGCGCAAGCTCTTAAAGG 4140  
QY 4141 GGAAGGACAGCTCCGCGCCCGCACCGCTCTCAGGGAAGCTGACCGAGGCTTACAAAGCGCA 4200  
Db 4141 GGAAGGACAGCTCCGCGCCCGCACCGCTCTCAGGGAAGCTGACCGAGGCTTACAAAGCGCA 4200  
QY 4201 GGGCCCTGGGCGCCCTGAAGCTGAAGCGGCGCATGAGGGGCTGGTGGCCAGGTTGAAGA 4260  
Db 4201 GGGCCCTGGGCGCCCTGAAGCTGAAGCGGCGCATGAGGGGCTGGTGGCCAGGTTGAAGA 4260  
QY 4261 GGGCGGCGCTTCATCATGAGATCCCGCGAGAGGCTGCGGACACGCCCGAGCTGCGC 4320  
Db 4261 GGGCGGCGCTTCATCATGAGATCCCGCGAGAGGCTGCGGACACGCCCGAGCTGCGC 4320  
QY 4321 CTTGGCCCGCGCGGCTCTCAAGAGGGCTTCATCAGCAGGGCACCCGCTCAAGTACGA 4380  
Db 4321 CTTGGCCCGCGCGGCTCTCAAGAGGGCTTCATCAGCAGGGCACCCGCTCAAGTACGA 4380  
QY 4381 CACCGGCGGCTCAACAATGAGCTCCAAAAGACGAGCTGCTCCATGAGGAGGCC 4440  
Db 4381 CACCGGCGGCTCAACAATGAGCTCCAAAAGACGAGCTGCTCCATGAGGAGGCC 4440  
QY 4441 CGGCGGAGAGTTCACACCCGTCACCCGCTGAGTGTATGCGGACCGCCGCGGACCTGGA 4500  
Db 4441 CGGCGGAGAGTTCACACCCGTCACCCGCTGAGTGTATGCGGACCGCCGCGGACCTGGA 4500  
QY 4501 AGTGCCTGCTACGAGAGAGCTGAAGAGCGGCGCAGGACCGCGCAGAGCTCGGAGG 4560  
Db 4501 AGTGCCTGCTACGAGAGAGCTGAAGAGCGGCGCAGGACCGCGCAGAGCTCGGAGG 4560  
QY 4561 CTGCATTGCGGCGGCGCGCGGCTGATTGTGCTGAGCTGGGTAAGCGCGGCAAGGCC 4620  
Db 4561 CTGCATTGCGGCGGCGCGCGGCTGATTGTGCTGAGCTGGGTAAGCGCGGCAAGGCC 4620  
QY 4621 CTTGACCTATAGAGACACGAGGACCCCTTTGCGGCGACCTTCCACAGAGTTCGCCGT 4680  
Db 4621 CTTGACCTATAGAGACACGAGGACCCCTTTGCGGCGACCTTCCACAGAGTTCGCCGT 4680  
QY 4681 GACCATGCGGAGGCCACGCGCGCTGCGAGAGGGCAGCTTTGTCTCAGCAAGGCAATC 4740  
Db 4681 GACCATGCGGAGGCCACGCGCGCTGCGAGAGGGCAGCTTTGTCTCAGCAAGGCAATC 4740

Db 4681 GACCATGCGGAGGCCACGCGCGCTGCGAGAGGGCAGCTTTGTCTCAGCAAGGCAATC 4740  
QY 4741 CCAGAGCCGAAAGCTGACGTGCACGCTCTGTAGATTCGCCAAGTCCCGCACAGACCGT 4800  
Db 4741 CCAGAGCCGAAAGCTGACGTGCACGCTCTGTAGATTCGCCAAGTCCCGCACAGACCGT 4800  
QY 4801 GCCGAGACCAACCCACACCCCATTTGCGCTTATGAGCACTTGCTTGGGGCTGATGG 4860  
Db 4801 GCCGAGACCAACCCACACCCCATTTGCGCTTATGAGCACTTGCTTGGGGCTGATGG 4860  
QY 4861 CGTGAAGCTGTATGAGACACCATCCCTGGCGTTGAGCCCACTCATATACCCGCGG 4920  
Db 4861 CGTGAAGCTGTATGAGACACCATCCCTGGCGTTGAGCCCACTCATATACCCGCGG 4920  
QY 4921 CATCCCTGTGACGAGCGCTGCTACTACTGCGCCGACACTTGGCCCCAACCCGAC 4980  
Db 4921 CATCCCTGTGACGAGCGCTGCTACTACTGCGCGGCTACCCGACAGCGGCGCTGGA 4980  
QY 4981 CTACCGGACCTGTACCCACCTTACTATCCGCGGCTACCCGACAGCGGCGCTGGA 5040  
Db 4981 CTACCGGACCTGTACCCACCTTACTATCCGCGGCTACCCGACAGCGGCGCTGGA 5040  
QY 5041 GAAACGGAGACATCATATGACTATGACTATGACTATGACTATGACTATGACTATGACTAT 5100  
Db 5041 GAAACGGAGACATCATATGACTATGACTATGACTATGACTATGACTATGACTATGACTAT 5100  
QY 5101 CACCGCATGCGCCAGAGCTGATGCTGAGGGGCTCTGCGCCGCGAGTCTCGCT 5160  
Db 5101 CACCGCATGCGCCAGAGCTGATGCTGAGGGGCTCTGCGCCGCGAGTCTCGCT 5160  
QY 5161 GGCATCTCAATACGTGCGGGTCCCGAGGATCATGCACTGTTCGAAGTCCACACTT 5220  
Db 5161 GGCATCTCAATACGTGCGGGTCCCGAGGATCATGCACTGTTCGAAGTCCACACTT 5220  
QY 5221 GCTGTGTGCTGCTGCGCGCGGACACAGAGGACCCGAGCGACCGCATGAGACCGCTTGGCTA 5280  
Db 5221 GCTGTGTGCTGCTGCGCGCGGACACAGAGGACCCGAGCGACCGCATGAGACCGCTTGGCTA 5280  
QY 5281 CTTCCCAACCGCGCCAGCCCTTCAAGCAGCGCCACAGAGTCCCATCTTCCGAGG 5340  
Db 5281 CTTCCCAACCGCGCCAGCCCTTCAAGCAGCGCCACAGAGTCCCATCTTCCGAGG 5340  
QY 5341 AGGTCAACACACTTGAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5400  
Db 5341 AGGTCAACACACTTGAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5400  
QY 5401 GATGAGAGAGGGGACCGGAGTGGGAGCGGAGAAAGTCCATCTTCACTTCCACGAC 5460  
Db 5401 GATGAGAGAGGGGACCGGAGTGGGAGCGGAGAAAGTCCATCTTCACTTCCACGAC 5460  
QY 5461 GGTGAGACGACCCCATCTGGAAGCTGTGTACAGAGGAGAGCGGACGAGCGGAG 5520  
Db 5461 GGTGAGAGACGACCCCATCTGGAAGCTGTGTACAGAGGAGAGCGGACGAGCGGAG 5520  
QY 5521 CAGCGCGGGGATGGGAGACAGACGCGGCGCTCCCATCTCCATCCATGCGCACAGCA 5580  
Db 5521 CAGCGCGGGGATGGGAGACAGACGCGGCGCTCCCATCTCCATCCATGCGCACAGCA 5580  
QY 5581 CTGCGCCATCTTCCCTCGGACCCAGAGTCCCTCAGAGAGAGACCCAGTGTCTTCAAA 5640  
Db 5581 CTGCGCCATCTTCCCTCGGACCCAGAGTCCCTCAGAGAGAGACCCAGTGTGTCAAA 5640  
QY 5641 CACAGGCAATGAAGGATATCATCACGCTGTGAGAGCCAGAGCCACGATCTTGAAGTTC 5700  
Db 5641 CACAGGCAATGAAGGATATCATCACGCTGTGAGAGCCAGAGCCACGATCTTGAAGTTC 5700  
QY 5701 CACCTCAACTTCAACCGTTGCGCGAGCTGCAATTCACCTGCACTGCACTGCCCC 5760  
Db 5701 CACCTCAACTTCAACCGTTGCGCGAGCTGCAATTCACCTGCACTGCACTGCCCC 5760  
QY 5761 ACTGGGCGGACCTTCAATGGGGTTCACCTTACCTTCAATGAGAGCCGTTGTGCGCAA 5820  
Db 5761 ACTGGGCGGACCTTCAATGGGGTTCACCTTACCTTCAATGAGAGCCGTTGTGCGCAA 5820

QY 5821 GAGAGCCCCGGGTGCGCCGGGCAAGAGCGAGCCCGAGCAGACACCGGCGATGCTTCT 5880  
| | | | |  
Db 5821 GAGAGCCCCGGGTGCGCCGGGCAAGAGCGAGCCCGAGCAGACACCGGCGATGCTTCT 5880  
| | | | |  
QY 5881 CGGCAAGCCCCCAAGCCCGCTCGGAGCTGAGAGCCCGCTCTCTCCCGAGCAAGGAGCTCGA 5940  
| | | | |  
Db 5881 CGGCAAGCCCCCAAGCCCGCTCGGAGCTGAGAGCCCGCTCTCTCCCGAGCAAGGAGCTCGA 5940  
| | | | |  
QY 5941 GCGCCGGGCCCCCTAGTGCTCTCTGCTCTTGAGCGACAGCCATGCGCCCGACCCCTTGGA 6000  
| | | | |  
Db 5941 GCGCCGGGCCCCCTAGTGCTCTCTGCTCTTGAGCGACAGCCATGCGCCCGACCCCTTGGA 6000  
| | | | |  
QY 6001 GAACCTGCACTCACAGCGCAGAGCCCGAGCCCGCGCGCAGCTTGCGCTCGA 6060  
| | | | |  
Db 6001 GAACCTGCACTCACAGCGCAGAGCCCGAGCCCGCGCGCAGCTTGCGCTCGA 6060  
| | | | |  
QY 6061 CGCGCAACGGGAAAGACTCAAGTAAACCTTTTCATCAGAGAACTGGAATCTCGTTTC 6120  
| | | | |  
Db 6061 CGCGCAACGGGAAAGACTCAAGTAAACCTTTTCATCAGAGAACTGGAATCTCGTTTC 6120  
| | | | |  
QY 6121 TCTGGGTACCAAGGCGAGGCTACAGCCCCGAAAGGGGTGAGAGCCGTCAGCCCTTGAG 6180  
| | | | |  
Db 6121 TCTGGGTACCAAGGCGAGGCTACAGCCCCGAAAGGGGTGAGAGCCGTCAGCCCTTGAG 6180  
| | | | |  
QY 6181 CTGACCCAGTCTGACCGCAGCAAGAGGGGCTCCCGAAGCAGCTGGAAGAGCTCGACAAAG 6240  
| | | | |  
Db 6181 CTGACCCAGTCTGACCGCAGCAAGAGGGGCTCCCGAAGCAGCTGGAAGAGCTCGACAAAG 6240  
| | | | |  
QY 6241 CCACTGAGAGGGGAGCTGCGGCGCAGCAGAGCCCGGTGAACTTGAGCGGAGAGG 6300  
| | | | |  
Db 6241 CCACTGAGAGGGGAGCTGCGGCGCAGCAGAGCCCGGTGAACTTGAGCGGAGAGG 6300  
| | | | |  
QY 6301 CGGCGCACTTCCCAACCTGCGGCGCTGCTGAGAGCAGCCCTGCTGAGCCCGCTGCT 6360  
| | | | |  
Db 6301 CGGCGCACTTCCCAACCTGCGGCGCTGCTGAGAGCAGCCCTGCTGAGCCCGCTGCT 6360  
| | | | |  
QY 6361 CGAGACCGGCGCGAGGGGTCAAAAGTCAACAGCGGGGTGTCAACCTGAGCCAGCAATCAG 6420  
| | | | |  
Db 6361 CGAGACCGGCGCGAGGGGTCAAAAGTCAACAGCGGGGTGTCAACCTGAGCCAGCAATCAG 6420  
| | | | |  
QY 6421 TGAAGTCATCAACAAGAGCTTACACCGGCAACAACCAAGAGCTGAGCGACCCCTGCC 6480  
| | | | |  
Db 6421 TGAAGTCATCAACAAGAGCTTACACCGGCAACAACCAAGAGCTGAGCGACCCCTGCC 6480  
| | | | |  
QY 6481 CGGCCCCCTTACTCTCTTCTGAGGCGCAGCTGCGCGTCTGAGACTTCCGCGCCAGC 6540  
| | | | |  
Db 6481 CGGCCCCCTTACTCTCTTCTGAGGCGCAGCTGCGCGTCTGAGACTTCCGCGCCAGC 6540  
| | | | |  
QY 6541 CAGTGAACCTTACTCTCTGCGGCGCGGCAACAAGTGAACCGGCGCGGCTCGCCAGCAG 6600  
| | | | |  
Db 6541 CAGTGAACCTTACTCTCTGCGGCGCGGCAACAAGTGAACCGGCGCGGCTCGCCAGCAG 6600  
| | | | |  
QY 6601 CGAAGGGGCGAAGAGGTCTTCAAGGCAACAAGAGCTGAGTGTGAGTGTGAGGA 6660  
| | | | |  
Db 6601 CGAAGGGGCGAAGAGGTCTTCAAGGCAACAAGAGCTGAGTGTGAGTGTGAGGA 6660  
| | | | |  
QY 6661 CGGTAATTGAACTGTGTCCCAACGAGGGGCAAGAGCGAGGCACTCCGAGTGC 6720  
| | | | |  
Db 6661 CGGTAATTGAACTGTGTCCCAACGAGGGGCAAGAGCGAGGCACTCCGAGTGC 6720  
| | | | |  
QY 6721 TGTGTATCCGGCTGTGATCCGGGATGGGAAACAGACCGAGCCAGAGATGGGCTTCGA 6780  
| | | | |  
Db 6721 TGTGTATCCGGCTGTGATCCGGGATGGGAAACAGACCGAGCCAGAGATGGGCTTCGA 6780  
| | | | |  
QY 6781 GTTCTCAGGCAACAGAGCGAGCGCAGCTTCTTCAAGAGCTGAACGAGCAATC 6840  
| | | | |  
Db 6781 GTTCTCAGGCAACAGAGCGAGCGCAGCTTCTTCAAGAGCTGAACGAGCAATC 6840  
| | | | |  
QY 6841 CGGCAATGCTCAAGTCAAGAGCAAGAGATCAACAGAACTGAACCCAGACCGGAA 6900  
| | | | |  
Db 6841 CGGCAATGCTCAAGTCAAGAGCAAGAGATCAACAGAACTGAACCCAGACCGGAA 6900  
| | | | |

QY 6901 TGAAGCTGAATATACATATACAGCAGCCTGAGGAGAGAGATCTTAATATGCCCCGATCAC 6960  
| | | | |  
Db 6901 TGAAGCTGAATATACATATACAGCAGCCTGAGGAGAGATCTTAATATGCCCCGATCAC 6960  
| | | | |  
QY 6961 CGGAACAGGCTTATGACCTATAGAAAGCGAGGGGTGAGAAATATCCAGACCAATCAT 7020  
| | | | |  
Db 6961 CGGAACAGGCTTATGACCTATAGAAAGCGAGGGGTGAGAAATATCCAGACCAATCAT 7020  
| | | | |  
QY 7021 GGGGCTGAGAGGCAATATTAAGAAAGGCACTATGGGTAAATATGACAGTGGGAAAGTGC 7080  
| | | | |  
Db 7021 GGGGCTGAGAGGCAATATTAAGAAAGGCACTATGGGTAAATATGACAGTGGGAAAGTGC 7080  
| | | | |  
QY 7081 CCGCGCTGAGGCGCAATGCTTTTAACTCTGAAATGCGAGTGCCAGCTTGCGCGTGC 7140  
| | | | |  
Db 7081 CCGCGCTGAGGCGCAATGCTTTTAACTCTGAAATGCGAGTGCCAGCTTGCGCGTGC 7140  
| | | | |  
QY 7141 TATGCCCATTAACCGCTGCTGACGAGAGTGAACACACTGACCTGCGCAGGTGCGG 7200  
| | | | |  
Db 7141 TATGCCCATTAACCGCTGCTGACGAGAGTGAACACACTGACCTGCGCAGGTGCGG 7200  
| | | | |  
QY 7201 CGGGAAGGCGCAAGGTCTGAGGAGAGCCAGCAGCCGAAAGCCAAAGTCCCGGCGCGG 7260  
| | | | |  
Db 7201 CGGGAAGGCGCAAGGTCTGAGGAGAGCCAGCAGCCGAAAGCCAAAGTCCCGGCGCGG 7260  
| | | | |  
QY 7261 CTTGAGATCTGAGGAGCGGCGCACTCTGCTCTGAGTGAAGTCTGAGAGGAGACTGCA 7320  
| | | | |  
Db 7261 CTTGAGATCTGAGGAGCGGCGCACTCTGCTCTGAGTGAAGTCTGAGAGGAGACTGCA 7320  
| | | | |  
QY 7321 CGGCGGAGCGCGCTGACCAACCGCTGTGAGAGAGAGAGGCGCTGTCGCAAGTTCAC 7380  
| | | | |  
Db 7321 CGGCGGAGCGCGCTGACCAACCGCTGTGAGAGAGAGAGGCGCTGTCGCAAGTTCAC 7380  
| | | | |  
QY 7381 GCGATTCCTCTTAACCCCTGATCATGCGGCTGAGGCGGGGTGATGAGCTTCCCGACC 7440  
| | | | |  
Db 7381 GCGATTCCTCTTAACCCCTGATCATGCGGCTGAGGCGGGGTGATGAGCTTCCCGACC 7440  
| | | | |  
QY 7441 CCCACCGGAGCTCCCGCGGAGCAGCGGAGCCCTGCGTGAAGCCCGCAGCAGCTTGAGAG 7500  
| | | | |  
Db 7441 CCCACCGGAGCTCCCGCGGAGCAGCGGAGCCCTGCGTGAAGCCCGCAGCAGCTTGAGAG 7500  
| | | | |  
QY 7501 GAGGCCAAGCACTGCTGTGCTGCAAGTACAGAGCACTTCCGAGCAGCAGTGAATCAG 7560  
| | | | |  
Db 7501 GAGGCCAAGCACTGCTGTGCTGCAAGTACAGAGCACTTCCGAGCAGCAGTGAATCAG 7560  
| | | | |  
QY 7561 AACAAGGCGGGGGGGGGGGGGGGGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 7620  
| | | | |  
Db 7561 AACAAGGCGGGGGGGGGGGGGGGGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 7620  
| | | | |  
QY 7621 CGGCGCTGAGAGCGGAGCGGCGCTGCGGACTCCCGCAACAAGAAAGAGAGCCCTGAGTC 7680  
| | | | |  
Db 7621 CGGCGCTGAGAGCGGAGCGGCGCTGCGGACTCCCGCAACAAGAAAGAGAGCCCTGAGTC 7680  
| | | | |  
QY 7681 CGGCTGCGCTCATCATCTGTCTGCTGCAAGAGCCGAGCATCTTGGCTGTCTAAAGCTTT 7740  
| | | | |  
Db 7681 CGGCTGCGCTCATCATCTGTCTGCTGCAAGAGCCGAGCATCTTGGCTGTCTAAAGCTTT 7740  
| | | | |  
QY 7741 AACTAAGACTCCCGGCGCGGCGCTGAGGCTTGAAGACTTAACTCAAGGAGATGTTTACCTG 7800  
| | | | |  
Db 7741 AACTAAGACTCCCGGCGCGGCGCTGAGGCTTGAAGACTTAACTCAAGGAGATGTTTACCTG 7800  
| | | | |  
QY 7801 GTGCTCGGGAAGGAGAGGAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7860  
| | | | |  
Db 7801 GTGCTCGGGAAGGAGAGGAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7860  
| | | | |  
QY 7861 ACAAGAGCGGCGAGGCGGCGGAGGAGCCCAAGCAGAGTGAACAGCAGCTTCAACGCGCAC 7920  
| | | | |  
Db 7861 ACAAGAGCGGCGAGGCGGCGGAGGAGCCCAAGCAGAGTGAACAGCAGCTTCAACGCGCAC 7920  
| | | | |  
QY 7921 TGCCTCCCGGAAATGATTTGGAACCAAGTCTAACTGAGCTGAGAGCCCGCGCGCTT 7980  
| | | | |  
Db 7921 TGCCTCCCGGAAATGATTTGGAACCAAGTCTAACTGAGCTGAGAGCCCGCGCGCTT 7980  
| | | | |  
QY 7981 CCTCGGCTCCCATCCGCTTGAAGCAGTGTGAGCAGATGAGCAGAGCCCTGTTCAGCCC 8040  
| | | | |  
Db 7981 CCTCGGCTCCCATCCGCTTGAAGCAGTGTGAGCAGATGAGCAGAGCCCTGTTCAGCCC 8040  
| | | | |

Db	7971	CCCTCCGCTCCATCCCGCTTAGCGCTCTGAGACAGATGACGAGGCCCTGTCCAGCC	8030
Qy	8041	CCAAGTGGAGCTCGTTCGCGTCCCAACAATGCGCCCAAGCAAGATTTGCTGAAACCA	8100
Db	8031	CCAAGTGGAGCTCGTTCGCGTCCCAACAATGCGCCCAAGCAAGATTTGCTGAAACCA	8090
Qy	8101	GTCAGGCGCAGGTGGGCGGACAAAGGGCCAGAGTGGCGCTGTGGGGGAAACGATGTCGGA	8160
Db	8091	GTCAGGCGCAGGTGGGCGGACAAAGGGCCAGAGTGGCGCTGTGGGGGAAACGATGTCGGA	8150
Qy	8161	GGACTGAGCTGTTTTTTTCAACAATCGTGTGCGGACGCGTGGGAAAGGAAAGGACAGATGTA	8220
Db	8151	GGACTGAGCTGTTTTTTTCAACAATCGTGTGCGGACGCGTGGGAAAGGAAAGGACAGATGTA	8210
Qy	8221	AATGATGTGTTGTTTACAGGGTATATTTTGTATACCTTCAATGATTAATTCAATGATGTT	8280
Db	8211	AATGATGTGTTGTTTACAGGGTATATTTTGTATACCTTCAATGATTAATTCAATGATGTT	8270
Qy	8281	TTACGCAAGGAGGACCTTACCAGTATTAATCTGCTGTGCTGCTTTGATCTGTCTTACCG	8340
Db	8271	TTACGCAAGGAGGACCTTACCAGTATTAATCTGCTGTGCTGCTTTGATCTGTCTTACCG	8330
Qy	8341	TTCAAGAGGCGGTGTGACAGGCGGACAGAGTGGGACCCCATCACTGCGAGGACCAAGGGGCG	8400
Db	8331	TTCAAGAGGCGGTGTGACAGGCGGACAGAGTGGGACCCCATCACTGCGAGGACCAAGGGGCG	8390
Qy	8401	GGGAGCTGCTCGTACAGCGCCCGCTGTGCTCTCCCTCCCTTCCCTTGGGACAGATGAA	8460
Db	8391	GGGAGCTGCTCGTACAGCGCCCGCTGTGCTCTCCCTCCCTTCCCTTGGGACAGATGAA	8450
Qy	8461	TTGCATGCGTATTTCTGTGGCGCCCATATTTGGCAGGGGTGTGTATTTCTGTATTTACACA	8520
Db	8451	TTGCATGCGTATTTCTGTGGCGCCCATATTTGGCAGGGGTGTGTATTTCTGTATTTACACA	8510
Qy	8521	CGTGCTCTAATTAAAAAGGCAATTATCTCCAAAAAATTTTTTTTTTTTTTTTTTTTTTTTT	8571
Db	8511	CGTGCTCTAATTAAAAAGGCAATTATCTCCAAAAAATTTTTTTTTTTTTTTTTTTTTTTTT	8561
RESULT 2	AF125672		
LOCUS	AF125672		
DEFINITION	Homo sapiens silencing mediator of retinoic acid and thyroid hormone receptor extended isoform (SMRTE) mRNA, complete cds.		
ACCESSION	AF125672		
VERSION	AF125672.1		
KEYWORDS	GI:4559297		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 8686)		
TITLE	Park,E.J., Schreen,D.J., Yang,M., Li,H., Li,L. and Chen,J.D.		
REFERENCE	SMRTE, a silencing mediator for retinoid and thyroid hormone receptors-extended isoform that is more related to the nuclear receptor corepressor		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3519-3524 (1999)		
PUBMED	10097068		
FEATURES	2 (bases 1 to 8686)		
REFERENCE	Chen,J.D.		
AUTHORS	Direct Submission		
TITLE	Submitted (03-FEB-1999) Pharmacology and Molecular Toxicology,		
JOURNAL	University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655, USA		
FEATURES	Location/Qualifiers		
SOURCE	1..8686		
gene	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	1..8686		
	/gene="SMRTE"		

[illegible]



Db 336 GCCCAGCGCGAGGCCCTCCCTGCTGTCTGAGTTTCAGAGCCCGGAGATGACGGTCCCA 395  
QY 241 GGAGCTCCACCTTGGCGGACAGAGTCCCACTGATCTGAGCCCGAGGCTGGGAGATCAAGAT 300  
Db 396 GGAGCTCCACCTTGGCGGACAGAGTCCCACTGATCTGAGCCCGAGGCTGGGAGATCAAGAT 455  
QY 301 GGAGTTCAATTGAAAGCAGCGCCTCGGCTAGAGCTGAGCTGAGCCCTGCTGCGACC 360  
Db 456 GGAGTTCAATTGAAAGCAGCGCCTCGGCTAGAGCTGAGCTGAGCCCTGCTGCGACC 515  
QY 361 GTCAACCCCTGCTGCGCAACCGGCGAGCTGCGGAGTCTGAAAGCTTCAACAGACCGTAG 420  
Db 516 GTCAACCCCTGCTGCGCAACCGGCGAGCTGCGGAGTCTGAAAGCTTCAACAGACCGTAG 575  
QY 421 CTTGACGGGCAAGCTGGAACCGGCTGTCCCGCAGCCCGGCAACCTGACCTGAGCT 480  
Db 576 CTTGACGGGCAAGCTGGAACCGGCTGTCCCGCAGCCCGGCAACCTGAGCT 635  
QY 481 GGAGCTGTGCGCGCAACCGCTGTCCAAGAGAGAGCTGATCCAAGACATGAGACCGCTGGA 540  
Db 636 GGAGCTGTGCGCGCAACCGCTGTCCAAGAGAGAGCTGATCCAAGACATGAGACCGCTGGA 695  
QY 541 CCGAGGATCAACATGTGTAGAGCAGCAGATCTTTAAGCTGAGAGAGAGCAGCAGACT 600  
Db 696 CCGAGGATCAACATGTGTAGAGCAGCAGATCTTTAAGCTGAGAGAGAGCAGCAGACT 755  
QY 601 GGAGAGAGAGGCTGCCAAGCCGCGCGAGCCTGGAAGCCCGGTGTCAACCGCGCGCAATCGA 660  
Db 756 GGAGAGAGAGGCTGCCAAGCCGCGCGAGCCTGGAAGCCCGGTGTCAACCGCGCGCAATCGA 815  
QY 661 GTTGAAGCACCAGAGCTGTGTGAGATCTTACGACGAGAAACCGGAGAGAGGCTGAAGC 720  
Db 816 GTTGAAGCACCAGAGCTGTGTGAGATCTTACGACGAGAAACCGGAGAGAGGCTGAAGC 875  
QY 721 TGCACATCGGATTTCTGAAAGGCTGTGGGCGCCCAAGGTGAGCTGCGCTGTACAACAGCC 780  
Db 876 TGCACATCGGATTTCTGAAAGGCTGTGGGCGCCCAAGGTGAGCTGCGCTGTACAACAGCC 935  
QY 781 CTCGCAACCCCGGAGTATCATGAGAACTCAAAATTAACCGAGCATGCGGAGAGAGCT 840  
Db 936 CTCGCAACCCCGGAGTATCATGAGAACTCAAAATTAACCGAGCATGCGGAGAGAGCT 995  
QY 841 AATCTTGTATCTTCAAGAGAGAGATCAAGCTCGGAAACAATGAGACAGAGATTCTGCGCA 900  
Db 996 AATCTTGTATCTTCAAGAGAGAGATCAAGCTCGGAAACAATGAGACAGAGATTCTGCGCA 1055  
QY 901 GCGCTATGACCAAGCTCATGAGGCTTGGAAAAAAGGTGAGCGCATGGAACCAACCC 960  
Db 1056 GCGCTATGACCAAGCTCATGAGGCTTGGAAAAAAGGTGAGCGCATGGAACCAACCC 1115  
QY 961 GCGCGCGCGGCGCAAGAGAGAGAGTGGCGGATCTTACGAAAAACAAGTCCCTGAGAT 1020  
Db 1116 GCGCGCGCGGCGCAAGAGAGAGAGTGGCGGATCTTACGAAAAACAAGTCCCTGAGAT 1175  
QY 1021 CCGCAAGCAGCGGAGCTGCAAGAGCGCATGCAAGAGGTGGGCGAGCGGCGGAGTGG 1080  
Db 1176 CCGCAAGCAGCGGAGCTGCAAGAGCGCATGCAAGAGGTGGGCGAGCGGCGGAGTGG 1232  
QY 1081 GCTGTTCATGTGTGGCGCGCGCGAGCAGAGCAGAGGTGTCAAGATCATGATGAGCTTTC 1140  
Db 1233 GCTGTTCATGTGTGGCGCGCGCGAGCAGAGCAGAGGTGTCAAGATCATGATGAGCTTTC 1292  
QY 1141 AAGAGCAGAGGAACTTGGAGAGAGAGATGGCGCAGCTGGGCTGTGATCCCGCCATGTGTGA 1200  
Db 1293 AAGAGCAGAGGAACTTGGAGAGAGAGATGGCGCAGCTGGGCTGTGATCCCGCCATGTGTGA 1352  
QY 1201 CGAGCTGACAGCAGCGCATCAAGTTCAATCAACATGAAACGGGCTTATGGCCGACCCAT 1260  
Db 1353 CGAGCTGACAGCAGCGCATCAAGTTCAATCAACATGAAACGGGCTTATGGCCGACCCAT 1412  
QY 1261 GAAAGGTGTACAAGACCGCGAGGTCTATGAACATGTGAGAGTGAAGAGAGAGACTTT 1320  
Db 1413 GAAAGGTGTACAAGACCGCGAGGTCTATGAACATGTGAGAGTGAAGAGAGAGACTTT 1472

QY 1321 CCGGAGAAAGTTGATGACAGATCCCAAGAACTTTGGCTGTATGCATCATTTCTGAGAG 1380  
Db 1473 CCGGAGAAAGTTGATGACAGATCCCAAGAACTTTGGCTGTATGCATCATTTCTGAGAG 1532  
QY 1381 GAAAGCAGTGGCTGAGTGGCTGCTTATTAATTAATTAATTAAGAGATGAGAACTATTA 1440  
Db 1533 GAAAGCAGTGGCTGAGTGGCTGCTTATTAATTAATTAATTAAGAGATGAGAACTATTA 1592  
QY 1441 GAGCTGTGTGAGCGAGAGCTATGTGGGCGCGGCAAGAGCAGCAGCAACAGCAGCA 1500  
Db 1593 GAGCTGTGTGAGCGAGAGCTATGTGGGCGCGGCAAGAGCAGCAGCAACAGCAGCA 1652  
QY 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
Db 1653 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1712  
QY 1561 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
Db 1713 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1772  
QY 1621 CCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
Db 1773 CCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1832  
QY 1681 GAAAGAGGCTGTGAGCTTCCAAAGAGCGCAAACTGCGCAACAGCGAGAGAGAGAGAG 1740  
Db 1833 GAAAGAGGCTGTGAGCTTCCAAAGAGCGCAAACTGCGCAACAGCGAGAGAGAGAGAG 1892  
QY 1741 CCGCATCAACCCGCTCAATGAGCTTAATGAGGCGCAACAGCGAGAGAGAGAGAGAGAG 1800  
Db 1893 CCGCATCAACCCGCTCAATGAGGCTTAATGAGGCGCAACAGCGAGAGAGAGAGAGAGAG 1952  
QY 1801 GAGCGCGAGCTGGCTTCCATGAGAGCTGTAATGAGAGTTCTCGTGGACAGAGAGAGAG 1860  
Db 1953 GAGCGCGAGCTGGCTTCCATGAGAGCTGTAATGAGAGTTCTCGTGGACAGAGAGAGAG 2012  
QY 1861 GAGAAACAGCAAGAAAGAGCTCTCTGGAGACAGGCGCGCAACTGCGGCAATCGCGGAT 1920  
Db 2013 GAGAAACAGCAAGAAAGAGCTCTCTGGAGACAGGCGCGCAACTGCGGCAATCGCGGAT 2072  
QY 1921 GGTGGCTTCAAGAGCTGTGTGCGAGTGTAGAACTTCTAATTCACTTAAGAGAGAGCA 1980  
Db 2073 GGTGGCTTCAAGAGCTGTGTGCGAGTGTAGAACTTCTAATTCACTTAAGAGAGAGCA 2132  
QY 1981 GAACTTCAATGAGATTTTGGACGACGACAAAGCTGAAAGATGAGAGAGAGAGAGAGAG 2040  
Db 2133 GAACTTCAATGAGATTTTGGACGACGACAAAGCTGAAAGATGAGAGAGAGAGAGAGAG 2192  
QY 2041 GAGAAAGAGAGAGAGAGAGCGCGCGCGCGCGAGCGAGAGAGAGAGAGAGAGAGAGAG 2100  
Db 2193 GAGAAAGAGAGAGAGAGAGCGCGCGCGCGCGAGCGAGAGAGAGAGAGAGAGAGAGAG 2252  
QY 2101 GAGAGATGAGAGATGAGAGAGCTGCGGCGGTGAGCGGAAATGAGAGAGAGATGAGAGAG 2160  
Db 2253 GAGAGATGAGAGATGAGAGAGCGGTGAGCGGCGGTGAGCGGAAATGAGAGAGAGATGAGAGAG 2212  
QY 2161 GGTGTAAGCTTATCATGCTCTTGGGAAATGAGGTGCCAAGAGGAGAAATGCAATGAGCCCAAGC 2220  
Db 2313 GGTGTA-----AGC 2321  
QY 2221 CATGTCAACCAAGCTCAGAGACCGAGAGCATCCCTCTCTCAACATGAGGCGCGCA 2280  
Db 2322 CATGTCAACCAAGCTCAGAGACCGAGAGCATCCCTCTCTCAACATGAGGCGCGCA 2381  
QY 2281 GAGCAGAGGCGAGATGAGGCGCGAGCGCGAGCAGCAGCTTGGGCGCGAGCGGAGAGAGAG 2340  
Db 2382 GAGCAGAGGCGAGATGAGGCGCGAGCGCGAGCAGCAGCTTGGGCGCGAGCGGAGAGAGAG 2441  
QY 2341 AGGCGTCAACCCAGCAGCAGAGAGATCCCGGCGCGCGAGCTTGAAGCCCAAGCGGCGCTC 2400  
Db 2442 AGGCGTCAACCCAGCAGCAGAGAGATCCCGGCGCGCGAGCTTGAAGCCCAAGCGGCGCTC 2501



[illegible]

Db	3582	GCCCCCTGCCCCATGAGACCCCAAAAAGCTGGCACCCCTTACGCGGAGTGAAGCAGAGCAGCT	3641
Qy	3517	GTCCCAACGAGAGCCAGAGCTTGGGCCACCCGAGAGCTTGAGAGTGGCCCAAGCCCAAGAGGC	3576
Db	3642	GTCCCCACGGGGGCAAGGCTTGGGCTCACCGGAGAGCTTGGGGGTGGCCCAAGGCCAGGAGGC	3701
Qy	3577	GTCCGTGCTGAGAGAGGACAGGCTCTGGGGCTCAGTTCCGGGGCGAAGACATCAAAAGGAT	3636
Db	3702	GTCCGTGCTGAGAGAGGACAGCTCTGGGGCTCAGTTCCGGGGCGAAGACATCAAAAGGAT	3761
Qy	3637	TCCACAGACACGGAGTCCCTCGGACAGCGCCCATCATACCGGGCTCATCACCCAAG	3696
Db	3762	TCCACAGACACGGAGTCCCTCGGACAGCGCCCATCATACCGGGCTCATCACCCAAG	3821
Qy	3697	CAGCCAGCTGAAGTCTCTGTACAGGGGACCATCAACAGATCATCGGCGAGAGCAGCCC	3756
Db	3822	CAGCCAGCTGAAGTCTCTGTACAGGGGACCATCAACAGATCATCGGCGAGAGCAGCCC	3881
Qy	3757	GAGTCGCTTGAAGCCGCGGGCGGGAGGACAGCTGCGCCAAAGGGCCACGTCATCTACGAAG	3816
Db	3882	GAGTCGCTTGAAGCCGCGGGCGGGAGGACAGCTGCGCCAAAGGGCCACGTCATCTACGAAG	3941
Qy	3817	CAAGAAAGGGCACAAGCTTGTGCTTACAGAGGAGTGAAGTCTGTGACCCAGTCTCTCAAGAA	3876
Db	3942	CAAGAAAGGGCACAAGCTTGTGCTTACAGAGGAGTGAAGTCTGTGACCCAGTCTCTCAAGAA	4001
Qy	3877	GAGCGGACAGAGCAGCTCAGAGCCCCCGCATGAGCGGCGCGCCCAAGCGCACCTATGA	3936
Db	4002	GAGCGGACAGAGCAGCTCAGAGCCCCCGCATGAGCGGCGCGCCCAAGCGCACCTATGA	4061
Qy	3937	CATGATGAGAGGGCGCGGTGGGCGAGAGCCATCTCTTCAGCCAGCATCCGAAGTCTCATGGG	3996
Db	4062	CATGATGAGAGGGCGCGGTGGGCGAGAGCCATCTCTTCAGCCAGCATCCGAAGTCTCATGGG	4121
Qy	3997	CGGTGCGATCCCGCGCGGAGCGACAGCGCCCGACCAAGCTCAAAAGGAGGACCAATCCG	4056
Db	4122	CGGTGCGATCCCGCGCGGAGCGACAGCGCCCGACCAAGCTCAAAAGGAGGACCAATCCG	4181
Qy	4057	CGGGTCATATCACAAAGGAGATCCCTCGGCTCTACGTGAGGCGACAGAGGAGCTACCTGCG	4116
Db	4182	CGGGTCATATCACAAAGGAGATCCCTCGGCTCTACGTGAGGAGGCGACAGAGGAGCTACCTGCG	4241
Qy	4117	TGCGGAGGCCAAGCTCTTAAAGCGGAGGAGGACGCGCTTCGCGCCCAACCGCCCTACGAGAA	4176
Db	4242	TGCGGAGGCCAAGCTCTTAAAGCGGAGGAGGACGCGCTTCGCGCCCAACCGCCCTACGAGAA	4301
Qy	4177	CTGACCCGAGGCGCTTAAAGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTATGA	4236
Db	4302	CTGACCCGAGGCGCTTAAAGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTATGA	4361
Qy	4237	GGGCTGTGTGACCAAGGTGAAGGAGGCGGGCGGCTCTCATTCATGAGATCCCGCGGAGAA	4296
Db	4362	GGGCTGTGTGACCAAGGTGAAGGAGGCGGGCGGCTCTCATTCATGAGATCCCGCGGAGAA	4421
Qy	4297	GCTGCGGACACAGCGCCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCAAGAGGCTCTCATAC	4356
Db	4422	GCTGCGGACACAGCGCCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCAAGAGGCTCTCATAC	4481
Qy	4357	GCAAGGGACCCCGGCTCAATATGAGACACCGGCGCGCTTCACACATGGCTTCCAAAAGCAGAA	4416
Db	4482	GCAAGGGACCCCGGCTCAATATGAGACACCGGCGCGCTTCACACATGGCTTCCAAAAGCAGAA	4541
Qy	4417	CGTACGCTCCCTATCGGAGAGCCCGGCGCGGACGTTCCCAACCGGTCACCCGCTGAGATGT	4476
Db	4542	CGTACGCTCCCTATCGGAGAGCCCGGCGCGGACGTTCCCAACCGGTCACCCGCTGAGATGT	4601
Qy	4477	GATGAGCGGACGCGCCGGGCACTGGAACGTGCGCTTACAGAGAGAGAGCTTGAAGAGCGCGCC	4536
Db	4602	GATGAGCGGACGCGCCGGGCACTGGAACGTGCGCTTACAGAGAGAGAGCTTGAAGAGCGCGCC	4661
Qy	4537	AGGAGACCGGACGAGCTCGAGGAGGCTCAATTGCGCGCGCGCGCGCGCGCTCATTTGTGCTTGA	4596

Db 4662 AGGAGCCGACGAGCTCGGGGGGCTCCATTGCGCGGCGCCCGGTCAATTGCTCTGA 4721  
OY 4597 GCTGGGTAAGCCGCGGAGAGCCCTGACCTTAAGAACAAGGAGCAACCTTTGGCGG 4656  
Db 4722 GCTGGGTAAGCCGCGGAGAGCCCTGACCTTAAGAACAAGGAGCAACCTTTGGCGG 4781  
OY 4657 CCACTTCCACGAGGTTCCGCGCTGACCATGCGGAGGCCACGCGGCTTGAGAGAGG 4716  
Db 4782 CCACTTCCACGAGGTTCCGCGCTGACCATGCGGAGGCCACGCGGCTTGAGAGAGG 4841  
OY 4717 CAGCTTTTGGTCCAGCAAGGCATCCAGGACCGAAAGCTGACGTGACGCTTGAGAT 4776  
Db 4842 CAGCTTTTGGTCCAGCAAGGCATCCAGGACCGAAAGCTGACGTGACGCTTGAGAT 4901  
OY 4777 CGCCAAAGTCCCGGACAGGACCGGTGCCGAGCAACCCACACCTCATCGGCTATGA 4836  
Db 4902 CGCCAAAGTCCCGGACAGGACCGGTGCCGAGCAACCCACACCTCATCGGCTATGA 4961  
OY 4837 GCACTGTGCGGGCGGTGAGTGGCGTGAACCTGTATCGGAGCAGATCCGCTGGCCTT 4896  
Db 4962 GCACTGTGCGGGCGGTGAGTGGCGTGAACCTGTATCGGAGCAGATCCGCTGGCCTT 5021  
OY 4897 CGAACCCACTCTCATATCCCGCGGCATCCCTCTGAGAGCAGCGGCTGCTACTACTGCC 4956  
Db 5022 CGAACCCACTCTCATATCCCGCGGCATCCCTCTGAGAGCAGCGGCTGCTACTACTGCC 5081  
OY 4957 CGGACACTGGGCGGCGGACCCCACTAACCGGCACTGTACCCGACTCACTCATCCGCGG 5016  
Db 5082 CGGACACTGGGCGGCGGACCCCACTAACCGGCACTGTACCCGACTCACTCATCCGCGG 5141  
OY 5017 CTACCCCGACAGCGCGCGCTGAGAAACCGGACAGACCATCATATGATCACTCACTTC 5076  
Db 5142 CTACCCCGACAGCGCGCGCTGAGAAACCGGACAGACCATCATATGATCACTCACTTC 5201  
OY 5077 GCAAGCATGACCAACAACAGCGCCACCGCATAGGCCAGCGAGCTGATATGCTGAGGGG 5136  
Db 5202 GCAAGCATGACCAACAACAGCGCCACCGCATAGGCCAGCGAGCTGATATGCTGAGGGG 5261  
OY 5137 CCTCTGCGCGCGGAGTCCCTGCTGCACTCACTAAGCTGCGGGTCCCGAGGACATCAT 5196  
Db 5262 CCTCTGCGCGCGGAGTCCCTGCTGCACTCACTAAGCTGCGGGTCCCGAGGACATCAT 5321  
OY 5197 CGACCTGTCCCAAGTGCCACACCTGCTGTGCTGTGCCCCGACACAGGCAACCCGACG 5256  
Db 5322 CGACCTGTCCCAAGTGCCACACCTGCTGTGCTGTGCCCCGACACAGGCAACCCGACG 5381  
OY 5257 CACCGCATGAGACCGCTTGTCTTACTCCACCGCGGCCAGCCCTTGACAGCGGCCA 5316  
Db 5382 CACCGCATGAGACCGCTTGTCTTACTCCACCGCGGCCAGCCCTTGACAGCGGCCA 5441  
OY 5317 CAGCAGCTCCCACTTCCCGAGGAGTCCAACAATTGAAGAAACCAACCAACCAAGTC 5376  
Db 5442 CAGCAGCTCCCACTTCCCGAGGAGTCCAACAATTGAAGAAACCAACCAACCAAGTC 5501  
OY 5377 CTCTGTCCGAGCGGAGCGAGAACCGGATCGAGAGCGGAGACCGGGATCGGGAGGAGAAA 5436  
Db 5502 CTCTGTCCGAGCGGAGCGAGAACCGGATCGAGAGCGGAGACCGGGATCGGGAGGAGAAA 5561  
OY 5437 GTTCATCTTCAAGTCCACGAGCGGTGAGCAACGACCATCTGAGACCTTGTAACA 5496  
Db 5562 GTTCATCTTCAAGTCCACGAGCGGTGAGCAACGACCATCTGAGACCTTGTAACA 5621  
OY 5497 GGAAGCAACCGGAGAGGCGGAGCGCGCGGGGTGGGGGAGAGAGAGCGCGCCCGC 5556  
Db 5622 GGAAGCAACCGGAGAGGCGGAGCGCGCGGGGTGGGGGAGAGAGAGCGCGCCCGC 5681  
OY 5557 CTCCCACTCCATGACCACAGCACTCGCCATCTCCCTCGAGACCAGGATCCCTTCA 5616  
Db 5682 CTCCCACTCCATGACCACAGCACTCGCCATCTCCCTCGAGACCAGGATCCCTTCA 5741  
OY 5617 GGAAGCAACCGAGTGTCTTCAACAACAAGGCATGAAGGGTATCATCACTGCTGAGACC 5676  
Db 5742 GGAAGCAACCGAGTGTCTTCAACAACAAGGCATGAAGGGTATCATCACTGCTGAGACC 5801

OY 5677 CAGCAAGCCCAAGGATCTGAGGTCAACTTCACTCTCAACCCGTTGACCACAGCTGCAC 5736  
Db 5802 CAGCAAGCCCAAGGATCTGAGGTCAACTTCACTCTCAACCCGTTGACCACAGCTGCAC 5861  
OY 5737 ATTCCCACTTGCCACCACTGCGCACTGGGCGGACCTCTGATGGGTCTTACCTTACCT 5796  
Db 5862 ATTCCCACTTGCCACCACTGCGCACTGGGCGGACCTCTGATGGGTCTTACCTTACCT 5921  
OY 5797 CATGAGCCCGCTTTCGTGCGCAAGAGGCCCCCGGGTCCCGGCAAGAGGCGCCG 5856  
Db 5922 CATGAGCCCGCTTTCGTGCGCAAGAGGCCCCCGGGTCCCGGCAAGAGGCGCCG 5981  
OY 5857 AGCAGACACCGGACATGCTTCTCGCCAAAGCCCGACCGGCTCCGGGCTGAGGCGCG 5916  
Db 5982 AGCAGACACCGGACATGCTTCTCGCCAAAGCCCGACCGGCTCCGGGCTGAGGCGCG 6041  
OY 5917 CTCTCTCCCAAGCAAGGCTCGAGGCCCCCGGCTTATGTCCTCTGTCTTGCCACGCG 5976  
Db 6042 CTCTCTCCCAAGCAAGGCTCGAGGCCCCCGGCTTATGTCCTCTGTCTTGCCACGCG 6101  
OY 5977 CACCATGCGCGGACCCCTGCGAAGAACTTCGCACTCAACGACGCAAGCCGAGCCGCGC 6036  
Db 6102 CACCATGCGCGGACCCCTGCGAAGAACTTCGCACTCAACGACGCAAGCCGAGCCGCGC 6161  
OY 6037 GGGCCACTGCTCGGCTCGGACCCGACCGGAGAAAGACTCAAGTAAACCTTTTC 6096  
Db 6162 GGGCCACTGCTCGGCTCGGACCCGACCGGAGAAAGACTCAAGTAAACCTTTTC 6221  
OY 6097 CATCAGAACTTGAATCTCGTTCTCTGAGGTTACACGCGAGAGTTACAGCCCGAAGG 6156  
Db 6222 CATCAGAACTTGAATCTCGTTCTCTGAGGTTACACGCGAGAGTTACAGCCCGAAGG 6281  
OY 6157 GGTGAGGCCCTGCAAGCCCTGTGAGTCAACCACTGACCCCAAGCAAGAGGCTCCCA 6216  
Db 6282 GGTGAGGCCCTGCAAGCCCTGTGAGTCAACCACTGACCCCAAGCAAGAGGCTCCCA 6341  
OY 6217 GCACTTGAAGAGCTGACAAAGAGCACTGAGAGGGAGCTGCGGCCCAAGAGCAGG 6276  
Db 6342 GCACTTGAAGAGCTGACAAAGAGCACTGAGAGGGAGCTGCGGCCCAAGAGCAGG 6401  
OY 6277 CCCCGTGAAGCTTGGCGGAGAGCGGCCCACTCCCAACTTGGCGGCTGTAAGAG 6336  
Db 6402 CCCCGTGAAGCTTGGCGGAGAGCGGCCCACTCCCAACTTGGCGGCTGTAAGAG 6461  
OY 6337 CCAAGCTTGTCTCAAGCCGCTGCTCAAGACCGGCCCAAGGGGTCAAGAGCGGT 6396  
Db 6462 CCAAGCTTGTCTCAAGCCGCTGCTCAAGACCGGCCCAAGGGGTCAAGAGCGGT 6521  
OY 6397 GGTCACTTGGCCAGCATCATGATGAGGTCAACAGAGACTACACCGGACACCC 6456  
Db 6522 GGTCACTTGGCCAGCATCATGATGAGGTCAACAGAGACTACACCGGACACCC 6581  
OY 6457 ACAGCAGCTCAAGGCAACCCCTGCGGCCCTCTTACTCTTCTTGGGCGCAGCTGCC 6516  
Db 6582 ACAGCAGCTCAAGGCAACCCCTGCGGCCCTCTTACTCTTCTTGGGCGCAGCTGCC 6641  
OY 6517 CGTCTTGAACCTTCGCGGCCCAACGATGACCTTACTCTTCGCGGCCCGGACATATGTC 6576  
Db 6642 CGTCTTGAACCTTCGCGGCCCAACGATGACCTTACTCTTCGCGGCCCGGACATATGTC 6701  
OY 6577 CCGGCGCGTGGCTCCCGCAAGGAGAGGGGCAAGAGGTCTCGAGGCAAAAGAGC 6636  
Db 6702 CCGGCGCGTGGCTCCCGCAAGGAGAGGGGCAAGAGGTCTCTCAGAGCAAAAGAGC 6761  
OY 6637 GTTCGATCTTGGGTGTGTGAGACGATATTGAACCTGTGTCCCAACGAGAGGCAATGAC 6696  
Db 6762 GTTCGATCTTGGGTGTGTGAGACGATATTGAACCTGTGTCCCAACGAGAGGCAATGAC 6821  
OY 6697 GGAAGCAAGGCACTCCCGAGAGTGTGTGAACCCGCTGCTGTACCGGAGATGGGAAAGAGC 6756  
Db 6822 GGAAGCAAGGCACTCCCGAGAGTGTGTGAACCCGCTGCTGTACCGGAGATGGGAAAGAGC 6881

QY	6757	GGAGCCCCGAGGAAATGGGCTCCAAAGTCTCCAGGGAAACAACAGCGAGCCGCCAGCTTCTT	6818
Db	6882	GGAGCCCAAGAGAAATGGGCTCCAAATCTTCAGGCAACAACAGCCAGCCGCGCAAGCTTCTT	6941
QY	6817	CAGCAAGCTGACCCGAGAACAACTCCGCCATGGTCAAGTCCAAAGACCAAGAAAGCAAGATCAACA	6876
Db	6942	CAGCAAGCTGACCCGAGAACAACTCCGCCATGGTCAAGTCCAAAGAAAGCAAGATCAACA	7001
QY	6877	GAAGCTGAACCCCAACAACCCGAATGAGCTTGAAATACAAATATCAGCCAGCCTGGACCGA	6938
Db	7002	GAAGCTGAACCCCAACAACCCGAATGAGCTTGAAATACAAATATCAGCCAGCCTGGACCGA	7061
QY	6937	GATCTTCAATATATGCCCCGCTATCACCGGAAACAGGCTTATGACCTATAGAACCAAGCCGCT	6998
Db	7062	GATCTTCAATATATGCCCCGCTATCACCGGAAACAGGCTTATGACCTATAGAACCAAGCCGCT	7121
QY	6997	GCAGGAACATGCAAGCAACCAACATGGGGCTGGAGGCTCAATATTTGAAAGGACCTCATGG	7056
Db	7122	GCAGGAACATGCAAGCAACCAACATGGGGCTGGAGGCTCAATATTTGAAAGGACCTCATGG	7181
QY	7057	TAAATATGACCAAGTGGGAAGATCCCGCGCTCAGCGCCAAATGCTTTAAACCTTGAA	7116
Db	7182	TAAATATGACCAAGTGGGAAGATCCCGCGCTCAGCGCCAAATGCTTTAAACCTTGAA	7241
QY	7117	TGCCAGTGCAGAGCTGCCCCGCTGCTATATGCCATTAACCCGCTGTCAAGCAAGATGACCA	7176
Db	7242	TGCCAGTGCAGAGCTGCCCCGCTGCTATATGCCATTAACCCGCTGTCAAGCAAGATGACCA	7301
QY	7177	CACACTCACTCGGACAGGTGGCGGGGGAAGGACCAAGGCTCTGGGCAACCCAGACGCG	7238
Db	7302	CACACTCACTCGGACAGGTGGCGGGGGAAGGACCAAGGCTCTGGGCAACCCAGACGCG	7361
QY	7237	AAAAGCCAAGTCCCGGCGCCCGGGGCTTGAGACTTGGGGACCGGCAACCTCTGTCTCTC	7298
Db	7362	AAAAGCCAAGTCCCGGCGCCCGGGGCTTGAGACTTGGGGACCGGCAACCTCTGTCTCTC	7421
QY	7287	AGTGCACCTCGGAGGGAGACTGCAACCCGCCGACCGCCCTCAACCAACCCGCTGTGGAGGA	7356
Db	7422	AGTGCACCTCGGAGGGAGACTGCAACCCGCCGACCGCCCTCAACCAACCCGCTGTGGAGGA	7481
QY	7357	CAGGCCCTCGTCCGAGGTTCCAGGCTCAAGCCATTCCTCCCAACACCCCGATCATAGCCGCTCA	7416
Db	7482	CAGGCCCTCGTCCGAGGTTCCAGGCTCAAGCCATTCCTCCCAACACCCCGATCATAGCCGCTCA	7541
QY	7417	GGCGGGTGTCAATGAGCTTCCCAACCCCAACCGGACTCCGCCCGGCAAGCGGAGCCCTGCG	7476
Db	7542	GGCGGGTGTCAATGAGCTTCCCAACCCCAACCGGACTCCGCCCGGCAAGCGGAGCCCTGCG	7601
QY	7477	TGGCCCCCAACAACCGCTGGAGACGAGAGACCCCAAGCACTGCTTGTCTGCAAGTACGAGAC	7538
Db	7602	TGGCCCCCAACAACCGCTGGAGACGAGAGACCCCAAGCACTGCTTGTCTGCAAGTACGAGAC	7661
QY	7537	ACTCTTCCGACAGCCAGTCACTCAAGAACAGGGCGGGGGGGGGCGGGCTGEBRSPERLUG	7598
Db	7662	ACTCTTCCGACAGCCAGTCACTCAAGAACAGGGCGGGGGGGGGCGGGCTGEBRSPERLUG	7711
QY	7597	TCAAGTCCCAAGAGCAACAGGAAACGGCCCTGACAGAAACGGGAGCGGCTGCGCACTCCCC	7656
Db	7712	TCAAGTCCCAAGAGCAACAGGAAACGGCCCTGACAGAAACGGGAGCGGCTGCGCACTCCCC	7771
QY	7657	AACCAAGGAAGAGCCCTCTGAGTCCGCTTGCGCTCATCTCATCTGTCCGATCAGAGCCG	7716
Db	7772	AACCAAGGAAGAGCCCTCTGAGTCCGCTTGCGCTCATCTCATCTGTCCGATCAGAGCCG	7831
QY	7717	GCATTCCTTGCCTGTCTAAAGCTTAACTAAAGACTTCGCCCCCGGGCTGCGCTGTGACGA	7776
Db	7832	GCATTCCTTGCCTGTCTAAAGCTTAACTAAAGACTTCGCCCCCGGGCTGCGCTGTGACGA	7891
QY	7777	CCTTACTCAGGGGAATTTACTCTGGTGTCCGGAAGGGAAGGGGGAAGGGGCGGGGGAGGG	7836
Db	7892	CCTTACTCAGGGGAATTTACTCTGGTGTCCGGAAGGGAAGGGGGAAGGGGCGGGGGAGGG	7951
QY	7837	GCACGGCAGGCGTGTGGCAGCCACAACAAGGCGGCAAGGCGGCTCAGAGGACCCAAAGCAG	7896

Db	7952	GCACGGCAGGCGTGTGGACCCACACACAGCGGCGGGCGGGACCCCAAGCAG	80111
Qy	7897	GATGACCACGACCTTCCACGCGCACTGCGCTCCCGGAATGATTTTGAACCAAGTCTAA	7956
Db	8012	GATGACCACGACCTTCCACGCGCACTGCGCTCCCGGAATGATTTTGAACCAAGTCTAA	8071
Qy	7957	CTGAGCTTGGACGCCCCCGCGCTTCCCTCCGCTTCCATTCGCCGTAGCGCTTGGACAG	8016
Db	8072	CTGAGCTTGGACGCCCCCGCGCTTCCCTCCGCTTCCATTCGCCGTAGCGCTTGGACAG	8133
Qy	8017	ATGACGCGAGGCGCTGTCAGCGCCCGAGTGGGCTCGTTCCGGTCCCGACAGCTGCCCCA	8076
Db	8132	ATGACGCGAGGCGCTGTCAGCGCCCGAGTGGGCTCGTTCCGGTCCCGACAGCTGCCCCA	8191
Qy	8077	GCCACAGAGATTTGCTGGAACCAAGTCAGGCGCAGGTGGCGGACAAAAGGCGCAGTGGC	8138
Db	8192	GCCACAGAGATTTGCTGGAACCAAGTCAGGCGCAGGTGGCGGACAAAAGGCGCAGTGGC	8251
Qy	8137	GCTGGGGGGGAAACGAGATCTCCGAGAGCTGACTGTTTTTTTTCACACATCGTTCCGCGAG	8196
Db	8252	GCTGGGGGGGAAACGAGATCTCCGAGAGCTGACTGTTTTTTTTCACACATCGTTCCGCGAG	8311
Qy	8197	CGGTGGGAAGGAAGCGAGATGTAATGATGTTGTTTACAGGGATATTTTGTATAC	8256
Db	8312	CGGTGGGAAGGAAGCGAGATGTAATGATGTTGTTTACAGGGATATTTTGTATAC	8371
Qy	8257	CTTCAATGAATTAATTCAGATGTTTTTACGAAAGAGACTTACCCAGTATTAATCTGCTGC	8316
Db	8372	CTTCAATGAATTAATTCAGATGTTTTTACGAAAGAGACTTACCCAGTATTAATCTGCTGC	8433
Qy	8317	TGTGCTTTTGAATCTCTGCTTACCGTTCAAGAGCGGTGTGCAAGCGCTGATACC	8376
Db	8432	TGTGCTTTTGAATCTCTGCTTACCGTTCAAGAGCGGTGTGCAAGCGCTGATACC	8491
Qy	8377	CATCACTCGCAGACCAAGGGGGGGGAGCTGCTGTACAGCGCCCGCTGATCTCCCTC	8436
Db	8492	CATCACTCGCAGACCAAGGGGGGGGAGCTGCTGTACAGCGCCCGCTGATCTCCCTC	8551
Qy	8437	CCTCCCTTCTCTTGGGACAGATGAATTCGATTCGATTTCTGTGCGCGCATTTTGGCAGGG	8496
Db	8552	CCTCCCTTCTCTTGGGACAGATGAATTCGATTCGATTTCTGTGCGCGCATTTTGGCAGGG	8611
Qy	8497	TGTGTGATTTCTGTCAATTACACAGCTCGTTCTAATTTAAAAAGCGAATTATCTCAAAA	8556
Db	8612	TGTGTGATTTCTGTCAATTACACAGCTCGTTCTAATTTAAAAAGCGAATTATCTCAAAA	8671
Qy	8557	AAAAAAAAAAAAAAAA 8571	
Db	8672	AAAAAAAAAAAAAAAA 8686	
RESULT 3			
LOCUS	AR447713	9053 bp	DNA
DEFINITION	Sequence 306 from patent US 6673549.	linear	PAT 20-FEB-2004
ACCESSION	AR447713		
VERSION	AR447713.1	GI:42676037	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 9053)		
AUTHORS	Furness,L.M. and Buchbinder,J.L.		
TITLE	Genes expressed in C3A liver cell cultures treated with steroids		
JOURNAL	Patent: US 6673549-A-306 06-UN-2004;		
FEATURES	Incyte Corporation; Palo Alto, CA		
source	1..9053		
ORIGIN	/organism="Unknown"		
	/mol_type="genomic DNA"		

Query Match 96.1%; Score 8226.8; DB 6; Length 9053;  
Best Local Similarity 98.3%; Pred. No. 0;  
Matches 8451; Conservative 2; Mismatches 42; Indels 103; Gaps 10;

QY 1 CATGTGGGCTCCACACAGCTTGTGGACAGACGTGAGGGCCACTGAGCCCCGCTAACCC 60  
DB 519 CATGTGGGATCCACACAGCCTGTGGACAGACGTGAGGGCCACTGAGCCCCGCTAACCC 578  
QY 61 GCGCCACAGCCTTTCCTACCCAGATGCGCCCGGACGACACCGACCTGGGCTCCCT 120  
DB 579 GCGCCACAGCCTTTCCTACCCAGATGCGCCCGGACGACACCGACCTGGGCTCCCT 638  
QY 121 GGAGTACACAGACCACTCCCGGACTATGCTCCCACTGCGCGGGCTCCATCATCCA 180  
DB 639 GGAGTACACAGACCACTCCCGGACTATGCTCCCACTGCTCCGCGGCTCATCATCCA 698  
QY 181 GCGCCACGCGCGGAGGCGCTCCCTGCTGTGAGTTCCAGCCCGGAAATGACGGTCCCA 240  
DB 699 GCGCCACGCGCGGAGGCGCTCCCTGCTGTGAGTTCCAGCCCGGAAATGACGGTCCCA 758  
QY 241 GGAGCTCCACCTGCGGCGGACGTCCTCATCTGCTGCGGAGCTGGGAAATCAGAGAT 300  
DB 759 GGAGCTCCACCTGCGGCGGACGTCCTCATCTGCTGCGGAGGATCAGAGAT 818  
QY 301 GGAGTTCATTGAAAGGAGCGCCTCGGCTTGAAGCTGCTGACCCCTGCTGCGACC 360  
DB 819 GGAGTTCATTGAAAGGAGCGCCTCGGCTTGAAGCTGCTGACCCCTGCTGCGACC 878  
QY 361 GTTACCCCTGTGCGGACGCGGCGGCTGCGGATGTGAAGCCTGACCAAGACCGTAG 420  
DB 879 GTTACCCCTGTGCGGACGCGGCGGCTGCGGATGTGAAGCCTGACCAAGACCGTAG 938  
QY 421 CCTGACGGGCAAGCTGGAACCGGTGTCTCCCGGACCGCGGACACTGACCTTGAAGCT 480  
DB 939 CCTGACGGGCAAGCTGGAACCGGTGTCTCCCGGACCGCGGACACTGACCTTGAAGCT 998  
QY 481 GGAGCTGTGCGGCGCACGGCTGTCCAAAGAGAGAGCTGATCCAGAACATGGAACCGGTGA 540  
DB 999 GGAGCTGTGCGGCGCACGGCTGTCCAAAGAGAGAGCTGATCCAGAACATGGAACCGGTGA 1058  
QY 541 CCGAGAGATCAACATGATGAGAGCAGACGATCTTAAAGCTGAAGAGAGACGACAGCT 600  
DB 1059 CCGAGAGATCAACATGATGAGAGCAGACGATCTTAAAGCTGAAGAGAGACGACAGCT 1118  
QY 601 GGAAGAGAGAGGCTGCCAAGCGCGCCGAGGCTGAAGACCGGTGTACCGCGCCCATCCA 660  
DB 1119 GGAAGAGAGAGGCTGCCAAGCGCGCCGAGGCTGAAGACCGGTGTACCGCGCCCATCCA 1178  
QY 661 GTTGAAGCACCGGAGCTGTGAGATCAATCTTACGAGAGACCGGAAAGGCTGAAGC 720  
DB 1179 GTTGAAGCACCGGAGCTGTGAGATCAATCTTACGAGAGACCGGAAAGGCTGAAGC 1238  
QY 721 TGCACATCGGATTTCTGAAAGGCTGAGGGGCCCAAGTGAAGCTCCCTGTACACAGGCC 780  
DB 1239 TGCACATCGGATTTCTGAAAGGCTGAGGGGCCCAAGTGAAGCTCCCTGTACACAGGCC 1298  
QY 781 CTTCGACACCCCGGAGTATCATGAGAACATCAAAATTAACACAGGCGATGCGAAAGACT 840  
DB 1299 CTTCGACACCCCGGAGTATCATGAGAACATCAAAATTAACACAGGCGATGCGAAAGACT 1358  
QY 841 AATCTTGTATCTTCAAGAGAGGAAATCAAGCTCGGAAACAAATGGAACGAAGTTCTGGCA 900  
DB 1359 AATCTTGTATCTTCAAGAGAGGAAATCAAGCTCGGAAACAAATGGAACGAAGTTCTGGCA 1418  
QY 901 GCGCTATGACCAAGCTATGAGAGGCTTGGAAAAAAGTGAAGCGATCGAAAAACAACC 960  
DB 1419 GCGCTATGACCAAGCTATGAGAGGCTTGGAAAAAAGTGAAGCGATCGAAAAACAACC 1478  
QY 961 GCGCGCGGCGGCGCAAGAGAGCAAGGTGCGGAGTACTACGAAAGCAAGTTCCCTTGAGAT 1020  
DB 1479 GCGCGCGGCGGCGCAAGAGAGCAAGGTGCGGAGTACTACGAAAGCAAGTTCCCTTGAGAT 1538  
QY 1021 CCGCAAGCAGCGGAGCTGACGAGAGGCAATGACAGAGAGGCTGGGCAAGCGGCGAGTGG 1080

DB 1539 CCGCAAGCAGCGGAGCTGACGAGAGCGCATGACAGACGAGGTGGCCAGCGGCGAGTGG 1598  
QY 1081 GCTGTCCATGTGGCCCGCCCGGACGACACGAGGTGTACAGATCATGATGCGCTTC 1140  
DB 1599 GCTGTCCATGTGGCCCGCCCGGACGACACGAGGTGTACAGATCATGATGCGCTTC 1658  
QY 1141 AGAGCAGAGAAACCTTGAGAGAGAGATGCGCCAGCTGCGCGGTATCCCGCCATGCTGTA 1200  
DB 1659 AGAGCAGAGAAACCTTGAGAGAGAGATGCGCCAGCTGCGCGGTATCCCGCCATGCTGTA 1718  
QY 1201 CGACGCTGACACGACGACGATCAAGTTCAATCAATGAAACGAGCTTATGCGACCCCAT 1260  
DB 1719 CGACGCTGACACGACGACGATCAAGTTCAATCAATGAAACGAGCTTATGCGACCCCAT 1778  
QY 1261 GAAAGTGTACAAAGACCCGCGGCTGTGAACATGTGAGTGAAGACAGAGAAAGACCTT 1320  
DB 1779 GAAAGTGTACAAAGACCCGCGGCTGTGAACATGTGAGTGAAGACAGAGAAAGACCTT 1838  
QY 1321 CCGGAGAAAGTTGATGACAGATCCCAAGAACTTGGCTGATGCAATCTTCTGGAGAG 1380  
DB 1839 CCGGAGAAAGTTGATGACAGATCCCAAGAACTTGGCTGATGCAATCTTCTGGAGAG 1898  
QY 1381 GAAAGCAGTGTGAGTGTGCTCTTATTACTTACTGACTAAGAGAAATGAGAACTATAA 1440  
DB 1899 GAAAGCAGTGTGAGTGTGCTCTTATTACTTACTGACTAAGAGAAATGAGAACTATAA 1958  
QY 1441 GAGCTGTGAGACGAGGCTATGTGCGCGCGGACAGACAGCAGACACAA--CAGCA 1497  
DB 1959 GAGCTGTGAGACGAGGCTATGTGCGCGCGGACAGACAGCAGACACAA--CAGCA 2018  
QY 1498 GCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1557  
DB 2019 GCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2078  
QY 1558 GAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGA 1617  
DB 2079 GAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGA 2138  
QY 1618 GAAACGACAAAGAGAACTCTCTCAAGAGAGAGACAGACGACCTCAGGGAGAGACAA 1677  
DB 2139 GAAACGACAAAGAGAACTCTCTCAAGAGAGAGAGACAGACGACCTCAGGGAGAGACAA 2198  
QY 1678 CAGAGAGAGGCTGTGCGCTCCAAAGCGCGCAAAATGCGCAACAGCAGGAGAGAGCA 1737  
DB 2199 CAGAGAGAGGCTGTGCGCTCCAAAGCGCGCAAAATGCGCAACAGCAGGAGAGAGCA 2258  
QY 1738 AGGCGCATCAACCGCTCAATGAGCTAATGAGGCAACAGCAGAGAGGCAATCAACCCCA 1797  
DB 2259 AGGCGCATCAACCGCTCAATGAGCTAATGAGGCAACAGCAGAGAGGCAATCAACCCCA 2318  
QY 1798 GCAAGAGCGCGAGCTGCTCCATGAGAGCTTAATGAGAGTTCGCTGACACAGAGAGA 1857  
DB 2319 GCAAGAGCGCGAGCTGCTCCATGAGAGCTTAATGAGAGTTCGCTGACACAGAGAGA 2378  
QY 1858 AATGAGAAACAGCAAGAAAGTGTCTTGAAGACAGGCGCAACTGTGCGGCTATCCCGG 1917  
DB 2379 AATGAGAAACAGCAAGAAAGTGTCTTGAAGACAGGCGCAACTGTGCGGCTATCCCGG 2438  
QY 1918 GATGTGGGCTCAAGACTGTGTGAGGTGAAGAACTTCACTTCAATCAAGAGAGAG 1977  
DB 2439 GATGTGGGCTCAAGACTGTGTGAGGTGAAGAACTTCACTTCAATCAAGAGAGAG 2498  
QY 1978 GCAAGAACTCGATGAGATTTTGACAGACACAAAGCTGAAGATGAGAGAGAGAGAG 2037  
DB 2499 GCAAGAACTCGATGAGATTTTGACAGACACAAAGCTGAAGATGAGAGAGAGAGAGAG 2558  
QY 2038 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2096  
DB 2559 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2618  
QY 2097 TGGTGAAGATGAGAGATGAGAGAGCTCGGCGGTGAGCGGAAATGAGAGAGAGATGTGA 2156

Db 2619 TGGTGAAGATGAGAGATGAGAGCGTCGGGCGGTGAGCGGAAATGAGAGAGAGATGCTGG 2678  
QY 2157 AGAGAGCTGAGAGCCTTACATGCTCTTGGAATGAGGTGCCAGAGGGGAATGCATGAGGCC 2216  
Db 2679 AGGAGGCTGA----- 2688  
QY 2217 CAGCCACTGTCAACACAGCTCAAGACCTGAGAGCATTCCTCTCTCTCACTGAGAGCCG 2276  
Db 2689 -AGCCACTGTCAACACAGCTCAAGACCTGAGAGCATTCCTCTCTCTCACTGAGAGCCG 2747  
QY 2277 CCAAGGACACAGGGGAGAAATGGGCCCCAGCCCCCAGCCACCTGTGGGCGCGAGC-GGGCGCA 2335  
Db 2748 CCAAGGACACAGGGGAGAAATGGGCCCCAGCCCCCAGCCACCTGTGGGCGCGAGGAGGCGCA 2807  
QY 2336 CCCCCAGGCCACCCACCCGACACCGAGAGACATCCGGGCCCCCAATTGAGCCACCCCG 2395  
Db 2808 CCCCCAGGCCACCCACCCGACACCGAGAGACATCCGGGCCCCCACTGAGGCCACCCCG 2867  
QY 2396 GCCTCTGAAGCCACCGGAGCCCTTACGCCCCCAACGACACCCCATGCTCTGTGACCT 2455  
Db 2868 GCCTCTGAAGCCACCGGAGCCCTTACGCCCCCAACGACACCCCATGCTCTGTGACCT 2927  
QY 2456 CCTCGTGTGTCCCAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2515  
Db 2928 CTTCTGTGTGTCCCAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2987  
QY 2516 GAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2575  
Db 2988 GAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3047  
QY 2576 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2635  
Db 3048 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3107  
QY 2636 GACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2695  
Db 3108 GACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3167  
QY 2696 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2755  
Db 3168 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3227  
QY 2756 GCTTACCTGAGTGCAGACGAGGTGATGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2815  
Db 3228 GCTTACCTGAGTGCAGACGAGGTGATGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 3287  
QY 2816 TCCCCAGAGGCCAGGCTCTCAACCCGAGTGGAGACCCCGAGGCCAATGCTCAACCCGAG 2875  
Db 3288 TCCCCAGAGGCCAGGCTCTCAACCCGAGTGGAGACCCCGAGGCCAATGCTCAACCCGAG 3347  
QY 2876 AAGCCACTGGAAGCTGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2935  
Db 3348 AAGCCACTGGAAGCTGAGAGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3404  
QY 2936 ACCAAAGTCAATGAGCCCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2994  
Db 3405 ACCAAAGTCAATGAGCCCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3464  
QY 2995 CCGACCGGACCGGCAAAACCTGACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3054  
Db 3465 CCGACCGGACCGGCAAAACCTGACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3524  
QY 3055 CCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3100  
Db 3525 CCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3584  
QY 3101 -----GCTTTCGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3150  
Db 3585 GTTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3644  
QY 3151 TTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3210  
Db 3645 TTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3704

QY 3211 GGAACCTTCAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3270  
Db 3705 GGAACCTTCAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3764  
QY 3271 CATGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3330  
Db 3765 CATGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3824  
QY 3331 CTCTGCAAGACCCCAAGAGCTCTGAGAGAGCAAAATAGGTGCATCTCCCAAGAAATGTC 3390  
Db 3825 CTCTGCAAGACCCCAAGAGCTCTGAGAGAGCAAAATAGGTGCATCTCCCAAGAAATGTC 3884  
QY 3391 GGTCCAGCTTCAGTCTCCGTACTCAGAGAGTCCAAAGGCCCGGTGGGCGCTGTCAAT 3450  
Db 3885 GGTCCAGCTTCAGTCTCCGTACTCAGAGAGTCCAAAGGCCCGGTGGGCGCTGTCAAT 3944  
QY 3451 GGGGCTGCGCCCTGCGCCATGAGAGCCCAAAAGGTGGCACCCCTTCAAGCGAGTGAAGAGGA 3510  
Db 3945 GGGGCTGCGCCCTGCGCCATGAGAGCCCAAAAGGTGGCACCCCTTCAAGCGAGTGAAGAGGA 4004  
QY 3511 GCAGCTGTCCCAAGGAGCCAGGCTGGGCGCACCGGAGAGCTTGGGAGTCCACAGCCCA 3570  
Db 4005 GCAGCTGTCCCAAGGAGCCAGGCTGGGCGCACCGGAGAGCTTGGGAGTCCACAGCCCA 4064  
QY 3571 GGAAGGCTTCGCTGTGAGAGGAGAGAGCTTGGGCTTCAAGTCCGGGCGGAAAGCATCA 3630  
Db 4065 GGAAGGCTTCGCTGTGAGAGGAGAGAGCTTGGGCTTCAAGTCCGGGCGGAAAGCATCA 4124  
QY 3631 AGGCATTCGAGCAACAGGAGTCCCTCGGACAGGCGCATCAATACCGGGCTTCATCAC 3690  
Db 4125 AGGCATTCGAGCAACAGGAGTCCCTCGGACAGGCGCATCAATACCGGGCTTCATCAC 4184  
QY 3691 CCACGAGCAGCAGAGCTGAGCTCTGTACAGGAGCAATCAACAGAGATCATGCGAGGA 3750  
Db 4185 CCACGAGCAGCAGAGCTGAGCTCTGTACAGGAGCAATCAACAGAGATCATGCGAGGA 4244  
QY 3751 CAGCCGAGTGGCTTGAACCGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3810  
Db 4245 CAGCCGAGTGGCTTGAACCGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4304  
QY 3811 CGAAGGCAAGAGGAGCAAGTCTTGTCTTATGAGAGGTGGCAATGTGTGAGCCCAATGCTC 3870  
Db 4305 CGAAGGCAAGAGGAGCAAGTCTTGTCTTATGAGAGGTGGCAATGTGTGAGCCCAATGCTC 4364  
QY 3871 CAAGGAGAGGCAAGAGAGAGTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3930  
Db 4365 CAAGGAGAGGCAAGAGAGAGTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4424  
QY 3931 CTATGACATGATGAGAGGCGCGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3990  
Db 4425 CTATGACATGATGAGAGGCGCGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4484  
QY 3991 CATGGGCGGTGCATCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4050  
Db 4485 CATGGGCGGTGCATCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4544  
QY 4051 CATCCGCGGATCATCAACAAGAGAGATCTTCAAGCTTCAAGTGAAGAGAGAGAGAGAGAG 4110  
Db 4545 CATCCGCGGATCATCAACAAGAGAGATCTTCAAGCTTCAAGTGAAGAGAGAGAGAGAGAG 4604  
QY 4111 CTTGCTGCGGAGAGAGAGCTTCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4170  
Db 4605 CTTGCTGCGGAGAGAGAGCTTCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4664  
QY 4171 AGGGAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4230  
Db 4665 AGGGAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4724  
QY 4231 CCAATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4290  
Db 4725 CCAATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4784



4291 CGAGGAGCTGGGACACCGCCGAGCTGGCCCGCGCGCTCAAGAGGCTC 4350  
4785 CGAGGAGCTGGGACACCGCCGAGCTGGCCCGCGCGCTCAAGAGGCTC 4844  
4351 CATCAAGAGGAGCCCGCTCAAGTACACACCGCGCTGTCACCTGGCTCAAAA 4410  
4845 CATCAAGAGGAGCCCGCTCAAGTACACACCGCGCTGTCACCTGGCTCAAAA 4904  
4411 GACGAGCGTACCTGCTCAAGGAGCCCGCGCGGAGGTTCCCAACCGGTGACCGCT 4470  
4905 GACGAGCGTACCTGCTCAAGGAGCCCGCGCGGAGGTTCCCAACCGGTGACCGCT 4964  
4471 GATGATGATGAGCGAGCCCGGAGCTGGAACGTGCTGTACAGAGAGCTGAGAG 4530  
4965 GATGATGATGAGCGAGCCCGGAGCTGGAACGTGCTGTACAGAGAGCTGAGAG 5024  
4531 CGGCGGAGGAGCGGAGAGCTCGGGGGGCTTCATTGCGCGCGCGCTCGTCAATTGT 4590  
5025 CGGCGGAGGAGCGGAGAGCTCGGGGGGCTTCATTGCGCGCGCGCTCGTCAATTGT 5084  
4591 GCTGAGCTGGGTAAAGCCCGGAGAGCCCGCTGACCTTATGAGAGACAGGAGGACCTT 4650  
5085 GCTGAGCTGGGTAAAGCCCGGAGAGCCCGCTTAACTTATGAGAGACAGGAGGACCTT 5144  
4651 TGCCGAGCACTCCACAGAGTTCCCGGTGACCATGCGGAGCCGACCGCGCTGCA 4710  
5145 TGCCGAGCACTCCACAGAGTTCCCGGTGACCATGCGGAGCCGACCGCGCTGCA 5204  
4711 GAGGAGGAGCTTTGCTCCAGAGGATCCACAGACCGAAAGCTGACGTGCGAGCTCG 4770  
5205 GAGGAGGAGCTTTGCTCCAGAGGATCCACAGACCGAAAGCTGACGTGCGAGCTCG 5264  
4771 TGAAGATGCGCAAGTCCCGGACAGACACGTGCTCGGAGACCAACCCATCTCGCC 4830  
5265 TGAAGATGCGCAAGTCCCGGACAGACACGTGCTCGGAGACCAACCCATCTCGCC 5324  
4831 CTATGAGCACTGCTTGGGGGGGTGAGTGGAGCTGTATCGAGGACATGCCCT 4890  
5325 CTATGAGCACTGCTTGGGGGGGTGAGTGGAGCTGTATCGAGGACATGCCCT 5384  
4891 GGCCTTCCGACCCGACCTCCATACCCCGGAGCATCCCTTGAAGAGAGCGCTGCTCAT 4950  
5385 GGCCTTCCGACCCGACCTCCATACCCCGGAGCATCCCTTGAAGAGAGCGCTGCTCAT 5444  
4951 CTTGCCCGGACACTGAGCCCGGACCCCACTTACCCTGACCTGTACCTCAT 5010  
5445 CTTGCCCGGACACTGAGCCCGGACCCCACTTACCCTGACCTGTACCTCAT 5504  
5011 CGCGGAGTACCCGAGACAGCGGAGGCTGAGAAACCGGAGACCATCATATGATCAT 5070  
5505 CGCGGAGTACCCGAGACAGCGGAGGCTGAGAAACCGGAGACCATCATATGATCAT 5564  
5071 CACTTCCGAGAGTACCAACAACAGGACACCGCATGAGCCGAGAGCTGATATGCT 5130  
5565 CACTTCCGAGAGTACCAACAACAGGACACCGCATGAGCCGAGAGCTGATATGCT 5624  
5131 GAGGGGCTTCTGCGCCCGGAGTCTTGTGCTGCACTCAACTAGCTGCGGTCCCGAG 5190  
5625 GAGGGGCTTCTGCGCCCGGAGTCTTGTGCTGCACTCAACTAGCTGCGGTCCCGAG 5684  
5191 CATCATTCGACTGTCCCAAGTGCACACCTGAGCTGTGCTGAGCCCGGACAGAGGAC 5250  
5685 CATCATTCGACTGTCCCAAGTGCACACCTGAGCTGTGCTGAGCCCGGACAGAGGAC 5744  
5251 CCAGGAGCAAGGAGAGAGCGCTTGCCTTACCTCCACAGCGCCGAGCCCTTCCAGAG 5310  
5745 CCAGGAGCAAGGAGAGAGCGCTTGCCTTACCTCCACAGCGCCGAGCCCTTCCAGAG 5804  
5311 CGGCGCAGAGAGTCCCGCTCTTCCAGAGAGGTCCAAACATTTGACAAACCAACAC 5370  
5805 CGGCGCAGAGAGTCCCGCTCTTCCAGAGAGGTCCAAACATTTGACAAACCAACAC 5864  
5371 CAGGTCTGTCCGAGCGGAGAGGAGACCGGAGTCCAGAGCGGAGCCGAGATCGGAGCG 5430

5865 CAGTCTCTGTCCGAGCGGAGAGAGACCGGAGTCCAGAGCGGAGCCGAGATCGGAGCG 5924  
5431 GAAAAAGTCCATCTTCAAGCTCCACCAAGAGTGAAGACAGCCATCTGAGAGCTGG 5490  
5925 GAAAAAGTCCATCTTCAAGCTCCACCAAGAGTGAAGACAGCCATCTGAGAGCTGG 5984  
5491 TACAGAGCAGAGAGCGGAGAGAGCGGAGAGCGGAGGAGTGGGGAGAGCAGAGCG 5550  
5985 TACAGAGCAGAGAGCGGAGAGAGCGGAGAGCGGAGGAGTGGGGAGAGCAGAGCG 6044  
5551 CCCCGCTTCCACTCCATGAGCCCAAGAGCTGAGCCATCTCCCTGAGACCAAGATGC 5610  
6045 CCCCGCTTCCACTCCATGAGCCCAAGAGCTGAGCCATCTCCCTGAGACCAAGATGC 6104  
5611 CTTCCAGAGAGAGCCAGTGTCTTCAACAACAGGATGAAAGGTATATACACCGCTGT 5670  
6105 CTTCCAGAGAGAGCCAGTGTCTTCAACAACAGGATGAAAGGTATATACACCGCTGT 6164  
5671 GAGGCCAGAGAGCCAGGATCTG-----AGGTCACTTCCACTCTCACTCGT 5721  
6165 GAGGCCAGAGAGCCAGGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6224  
5722 TGCGGAGCTGACATTTCCACCTTGCACCCGATGCGGAGGAGGAGGAGGAGGAGGAG 5781  
6225 TGCGGAGCTGACATTTCCACCTTGCACCCGATGCGGAGGAGGAGGAGGAGGAGGAG 6284  
5782 GGTCTACCTTACCTTATGAGAGCCGCTTGTGCTGCGGAGAGAGGAGGAGGAGGAGGAG 5841  
6285 GGTCTACCTTACCTTATGAGAGCCGCTTGTGCTGCGGAGAGAGGAGGAGGAGGAGGAG 6344  
5842 GCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5901  
6345 GCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6404  
5902 CGGAGTGAAGCCGCTCTCTCCCGAGAGAGGATCGAGAGCCCGGAGCCCTAGTGCCTC 5961  
6405 CGGAGTGAAGCCGCTCTCTCCCGAGAGAGGATCGAGAGCCCGGAGCCCTAGTGCCTC 6464  
5962 TGTCTTGGGACAGGACCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6021  
6465 TGTCTTGGGACAGGACCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6524  
6022 CAGCCCGGAGCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6081  
6525 CAGCCCGGAGCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6584  
6082 AAGTAAACCTTTTCCATCCAGAACTGAACTCCGTTCTTGGTTACACGAGGAGAG 6141  
6585 AAGTAAACCTTTTCCATCCAGAACTGAACTCCGTTCTTGGTTACACGAGGAGAG 6644  
6142 CTACAGCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6201  
6645 CTACAGCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6704  
6202 CAAAGGAGTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6261  
6705 CAAAGGAGTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6764  
6765 GCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6824  
6322 GCGGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6381  
6825 GCGGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6884  
6382 AGGTACACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6441  
6885 AGGTACACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6944  
6442 CAGCGGAGCAGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6501



Dh	6545	CA	CCCGG	CA	CCCA	CCCA	CAG	CA	AG	CT	CAG	GG	CA	CCCT	TG	CCG	CC	CC	CT	CT	ACT	CT	TCC	7004
Qy	6502	TG	GGG	CA	CG	TG	CC	CC	CT	GT	GA	CT	T	CG	CC	CA	CC	CA	GT	GA	CT	T	CCG	6561
Dh	7005	TG	GGG	CA	CG	TG	CC	CC	CT	GT	GA	CT	T	CG	CC	CA	CC	CA	GT	GA	CT	T	CCG	7064
Qy	6562	CC	CGG	CA	CA	TG	GT	CC	CC	GG	CC	GG	CT	CC	CC	CA	CA	GG	CA	GA	GT	CT	CC	6621
Dh	7065	CC	CGG	CA	CA	TG	GT	CC	CC	GG	CC	GG	CT	CC	CC	CA	CA	GG	CA	GA	GT	CT	CC	7124
Qy	6622	AGA	CC	CA	AA	CA	GA	CG	TG	CG	CT	TG	GG	GT	GT	GA	GA	CG	GT	AT	TA	GA	CC	6681
Dh	7125	AGA	CC	CA	AA	CA	GA	CG	TG	CG	CT	TG	GG	GT	GT	GA	GA	CG	GT	AT	TA	GA	CC	7184
Qy	6682	AC	CGA	GG	GG	CA	TGA	CG	GA	CC	GA	GG	CA	CT	CC	CG	AG	TG	CT	GT	GA	CC	CG	6741
Dh	7185	AC	CGA	GG	GG	CA	TGA	CG	GA	CC	GA	GG	CA	CT	CC	CG	AG	TG	CT	GT	GA	CC	CG	7244
Qy	6742	GG	AT	GG	GG	AA	CA	GA	CG	GA	CC	GA	GG	AT	GG	CT	CC	AA	GT	CT	CA	GG	CA	6801
Dh	7245	GG	AT	GG	GG	AA	CA	GA	CG	GA	CC	GA	GG	AT	GG	CT	CC	AA	GT	CT	CA	GG	CA	7304
Qy	6802	GC	CG	CA	GG	CC	CT	CT	T	CA	GA	CG	AG	CT	GA	CG	CA	CT	CC	CG	CA	TG	CT	6861
Dh	7305	GC	CG	CA	GG	CC	CT	CT	T	CA	GA	CG	AG	CT	GA	CG	CA	CT	CC	CG	CA	TG	CT	7364
Qy	6862	GCA	AG	AT	CA	CA	AA	AG	CT	GA	CA	CC	CA	CA	CC	GA	AT	GA	GC	CT	GA	TA	T	6921
Dh	7365	GCA	AG	AT	CA	CA	AA	AG	CT	GA	CA	CC	CA	CA	CC	GA	AT	GA	GC	CT	GA	TA	T	7424
Qy	6922	CC	AG	CT	GG	GA	CG	GA	AT	CT	CA	T	A	T	AT	AT	AT	AT	AT	AT	AT	AT	6981	
Dh	7425	CC	AG	CT	GG	GA	CG	GA	AT	CT	CA	T	A	T	AT	AT	AT	AT	AT	AT	AT	AT	7484	
Qy	6982	TAGA	AG	CC	AG	GG	CG	GT	GA	CG	AA	CA	T	AT	CG	AG	CA	CA	A	T	AT	AT	7041	
Dh	7485	TAGA	AG	CC	AG	GG	CG	GT	GA	CG	AA	CA	T	AT	CG	AG	CA	CA	A	T	AT	AT	7544	
Qy	7042	AAA	GC	AT	CA	TG	GG	TAA	T	A	T	GA	CA	T	GA	GG	AA	AG	T	CC	CG	CG	CT	7101
Dh	7545	AAA	GC	AT	CA	TG	GG	TAA	T	A	T	GA	CA	T	GA	GG	AA	AG	T	CC	CG	CG	CT	7604
Qy	7102	TTTT	AA	CC	CT	CT	GA	AT	GC	CA	GT	GC	CA	GT	GC	CA	TA	AC	CG	CT	GA		7161	
Dh	7605	TTTT	AA	CC	CT	CT	GA	AT	GC	CA	GT	GC	CA	GT	GC	CA	TA	AC	CG	CT	GA		7664	
Qy	7162	CG	GA	CG	AG	TGA	CA	CA	CT	CA	CT	CG	CA	CG	GT	GG	GG	GG	GG	GG	GG	GG	7221	
Dh	7665	CG	GA	CG	AG	TGA	CA	CA	CT	CA	CT	CG	CA	CG	GT	GG	GG	GG	GG	GG	GG	GG	7724	
Qy	7222	CAGA	CC	CA	GA	CG	AG	CG	AA	AG	CA	GT	CC	CG	CG	CC	CG	CG	CG	CG	CG	CG	7281	
Dh	7725	CAGA	CC	CA	GA	CG	AG	CG	AA	AG	CA	GT	CC	CG	CG	CC	CG	CG	CG	CG	CG	CG	7784	
Qy	7282	AC	CT	CT	GT	CT	CT	GA	GT	GA	CT	CG	AA	GG	AT	CT	GA	CG	CA	CT	CC	CT	7341	
Dh	7785	AC	CT	CT	GT	CT	CT	GA	GT	GA	CT	CG	AA	GG	AT	CT	GA	CG	CA	CT	CC	CT	7844	
Qy	7342	CC	CG	GT	GG	GA	GA	CA	GG	CC	CT	GT	CC	GA	GG	TT	CA	CG	CA	TT	CC	CT	7401	
Dh	7845	CC	CG	GT	GG	GA	GA	CA	GG	CC	CT	GT	CC	GA	GG	TT	CA	CG	CA	TT	CC	CT	7904	
Qy	7402	GAT	CAT	G	CG	CT	G	CA	G	G	G	GT	CAT	TG	G	CT	CC	CA	CC	CC	CA	CC	CG	7461
Dh	7905	GAT	CAT	G	CG	CT	G	CA	G	G	G	GT	CAT	TG	G	CT	CC	CA	CC	CC	CA	CC	CG	7964
Qy	7462	CAG	CG	G	G	C	C	T	G	G	C	C	C	C	C	C	CA	CA	G	C	T	G	7521	
Dh	7965	CAG	CG	G	G	C	C	T	G	G	C	C	C	C	C	C	CA	CA	G	C	T	G	8024	
Qy	7522	CT	CG	CA	G	T	A	CG	A	CA	CT	T	CC	GA	CAG	GA	AT	GA	CT	CA	GA		7581	
Dh	8025	CT	CG	CA	G	T	A	CG	A	CA	CT	T	CC	GA	CAG	GA	AT	GA	CT	CA	GA		8084	

[illegible]

RESULT 4			
AY965853			
LOCUS	8548 bp	mRNA	linear
AY965853			PRI 11-APR-2005

DEFINITION	Homo sapiens SMRT- $\tau$ mRNA, complete cds, alternatively spliced.
ACCESSION	AY965853
VERSION	AY965853.1 GI:62240097
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Bakarjota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
REFERENCE	1 (bases 1 to 8548)
AUTHORS	Chen,J.D. and Evans,R.M.
TITLE	A transcriptional co-repressor that interacts with nuclear hormone receptors
JOURNAL	Nature 377 (6548), 454-457 (1995)
PUBMED	7566127
REFERENCE	2 (bases 1 to 8548)
AUTHORS	Park,E.J., Schreier,D.J., Yang,M., Li,H., Li,L. and Chen,J.D.
TITLE	SMRT, a silencing mediator for retinoid and thyroid hormone receptors-extended isoform that is more related to the nuclear receptor corepressor
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3519-3524 (1999)
PUBMED	10097068
REFERENCE	3 (bases 1 to 8548)
AUTHORS	Li,C.-W., Johnson,D., Chen,L.-Y., Ghosh,J. and Chen,J.D.
TITLE	Regulation and binding of the steroids and xenobiotics receptor SXR/PXR by nuclear receptor corepressor SMRT
JOURNAL	Unpublished
PUBMED	4 (bases 1 to 8548)
REFERENCE	Chen,J.D.
AUTHORS	Direct Submission
TITLE	Submitted (17-MAR-2005) Pharmacology, UMDNJ-Robert Wood Johnson Medical School, 661 Hoes Lane, Piscataway, NJ 08901, USA
JOURNAL	Location/Qualifiers
PUBMED	1. 8548
REFERENCE	/organism="Homo sapiens"
AUTHORS	/mol_type="mRNA"
TITLE	/db_xref="taxon:9606"
JOURNAL	157. 7542
PUBMED	/note="silencing mediator of retinoid and thyroid hormone receptors extended isoform splice variant 1;
REFERENCE	transcriptional corepressor; SMRT; alternatively spliced"
AUTHORS	/codon_start=1
TITLE	/product="SMRT- $\tau$ "
JOURNAL	/protein_id="AA77219.1"
PUBMED	/db_xref="GI:62240098"
REFERENCE	/translation="MSGQPVAVQWTRATPRPHSLSYVQIATHTDVLGLEYQH
AUTHORS	HSRDVASHLSPGSIIOQRRLSLSEFQGNERSQELHPRSHLYLBLEKSEMEF
TITLE	LSKPRRLPLDPLLRPSPLATQGPASEDLTKDRSLTGKLEPVSPSPHTDEL
JOURNAL	ELVPPRLSKBELQNMNDYDRREITWEOQISLKKQQLLEBAKPPPEKVPSP
PUBMED	IESKHSLOIYIDENRKAABAHRLLEGQVVELPLYNOPDRQYHENTKINQAM
REFERENCE	RKLLIYFRRBHRAKQWOKPCQRYDQMEAMEKVEEITENPRPRAESKRYEYE
AUTHORS	KQPEIRKORBLDERKQVCGSGSLSMSPASHEVSEITDLSQENLEKQMRQIA
TITLE	VLPMLYDADQRIKFINNGLMADPMKYIKORQVNMWSEDEKETFRKFMQHPNF
JOURNAL	GIASLERKTVAEVLYYLITKKNYSILVRSYRSGKSGQQQQQQQQQQQQQQ
PUBMED	QMPRSQOEKDEKEKEKEKEKEKEKEKEKEKEKEKEDLKEKTDITSENDKEKAVAK
REFERENCE	GKRTANSQGRKGRITRSMANRANSEALITPOOASLAMELNESRMTSEETAKK
AUTHORS	GILHEGRNWSATAPWGSKTVCQCKNPFYNYKQWLDILQHKIKMEKRNARBK
TITLE	KAPPAASERBAFPVYDEBEMASGVYTNBEMVBEALTYVNTSDTISPSPTTA
JOURNAL	AADTQNGKRPATLIGADGPPGPPPPEDIPAPTESIPASATLAPLPBPAPPS
PUBMED	SPPVAPKEKEKETAAAPVEGEQKPAEELAVDGKAEKPEVSCTEBAEAGP
REFERENCE	AGCKDEAEATATRALKAKEGSGATTAASGAPODSDATCSADVEDBAEG
AUTHORS	DKNRLLSPRLITPTGDRANASPOKPLDKQKORAAPIPIQVTVKPPAPBAAO
TITLE	KLPGRPPCTSGLPFPVPPREVITKASPHAPDPSAPSYAPRPHPLPLGLDTPALVLR
JOURNAL	PLTISPPRLISSAKHPSVLEKQIGAIISQMSVQALVPSBNAKAVGVSTGLPLPM
PUBMED	DKKCAPFGSVQOEQSPRGQAPPELSQVPTAOEASVLRGAGVGVSTKGIIPS
REFERENCE	TVPSDSATITVRSITHTPADVLTKGTTIRLIGEDSPRLGRGSRALPKSGIIEG
AUTHORS	KKGHVLSYSGMSVTCSEKDESGSGPHEHTAPKRTYDMMEGRGRALISSAIGEL
TITLE	MRALIPRHSPIHLKEOHNGISTOGIPRAVYEAQEDYLIRREALIRBELGPPPP
JOURNAL	PSRDLTEAKTQALPLKPKRNEGVLATVYKAGRSINHPREBLNTEBELPLAPPL
PUBMED	KGSITQGTPLKTDGASTTGSKHDVRLVLSGKTPFPVPHLDVMAADRALERCY
REFERENCE	ESLSKRPSTASSSGSIARGAIVIVBELGKPSQSLTVEHPHAPAGHLPGSPVTM
AUTHORS	REPTRLQSGSLSSSKASQDRKLTSTPREIAKSPHSTVEBHHPIPSIYEHILRGVS

VDLYKSHIPLADPISIPNGIPLDAAAAYLPHNLANPPTYPLYPYLIRGPTPTA  
LENRQTIINDYITSQQMHNNTATAMAQRADMLRGLSPRSSLLANTAAERGLITDLSQ  
VPHLPVLVPTPTGTPATMDRLAYLPTAQPFSSRSSSPHLPKPTPTSS  
ERRDRDRDRDREREKSIITSTTVENAHPIWRPTBOSSGSGSSGSSSPRA  
SHHAROHPIPRQDLOORPSVLANHGMKIIITAVEBSKPTVSTGSSPVPA  
ATPPATHTPDLGTTLDGVYPTLMEPVLIRKEARVARVPRPADQHAFLAKPARSG  
LEPASPSSKSGSEPRPLVPPVSGATTAARTPAKUALPHNASPPDPAAPASDPRREKT  
QSKPFSIOBLEKSLIGYHSSYSPEVEPVSPSLTHDKLPHLELSB  
ELRKOPGPVUKLGEBAHLPHLRPLPESQPSPLQTAAGVGHORVWTLAQHISB  
ITODYRHHNPOOLAPLAPLAPLFPASCPVLDLRPPDPLVPPDHGAPARSPHS  
EGKRSPEPNKTSVIGGEGDIEPVSPRPMTEBGRSRAVYPLLYRDGOSTPSMG  
SKSPANTSGPPAPRPSKLTESNANMYSKQBIINKULNTNRNPEVNIQDPRGRTNM  
PATGTGLMTYRSQAVQERASTWMLGIALIRKALMGSGKAKVSGRPSSKASLAPG  
LASGRPPSVSVSHBGDNRRTPLTNRWEDRSPSAGSGTTPPEYNPLIWRLOAGVNAS  
PPPGPLPAGSGPLAGPHHAWDEBPKPLCSQYETLSDSE"

Query Match 93.9%; Score 8038.6; DB 8; Length 8548;  
Best Local Similarity 97.0%; Pred. No. 0;  
Matches 8334; Conservative 1; Mismatches 34; Indels 226; Gaps 5;

Qy	1	CATGTGGGCTCCACACAGCTTGTCACAGACGTGAGAGGCGCATGAGCCCGCTAACCC	60
Db	156	CATGTGGGCTCCACACAGCTTGTCACAGACGTGAGAGGCGCATGAGCCCGCTAACCC	215
Qy	61	GCCCCACAGCTTTCTTACCCAGTGCAGATCGCCCGGACGACACAGAGCTCGAGCTCCT	120
Db	216	GCCCCACAGCTTTCTTACCCAGTGCAGATCGCCCGGACGACACAGAGCTCGAGCTCCT	275
Qy	121	GGAGTACAGACACACTCCCGGACATATGCTCTCCACCTGTGCGCGGCTCATCATCA	180
Db	276	GGAGTACAGACACACTCCCGGACATATGCTCTCCACCTGTGCGCGGCTCATCATCA	335
Qy	181	GCCCCAGCGCGGAGGCGCTCTGCTGTGATGTTCCAGCCCGGAAATGAACGGTCCCA	240
Db	336	GCCCCAGCGCGGAGGCGCTCTGCTGTGATGTTCCAGCCCGGAAATGAACGGTCCCA	395
Qy	241	GGAGCTCCACCTGCGGCGACAGTCCCATCATCTGCGCGAGCTGGGAAATGACAGAT	300
Db	396	GGAGCTCCACCTGCGGCGACAGTCCCATCATCTGCGCGAGCTGGGAAATGACAGAT	455
Qy	301	GGAGTTCATTGAAAGAGACGCGCTTCGAGTGAAGCTGCTGACCCCTGCTGAGCC	360
Db	456	GGAGTTCATTGAAAGAGACGCGCTTCGAGTGAAGCTGCTGACCCCTGCTGAGCC	515
Qy	361	GTCAACCCCTGCTGCGACAGGCGGCGAGCTCGGGATTTGAAAGACTTACCAAGAACCTGAG	420
Db	516	GTCAACCCCTGCTGCGACAGGCGGCGAGCTCGGGATTTGAAAGACTTACCAAGAACCTGAG	575
Qy	421	CTGACGGGCAAGCTGGAACCGGATGCTCCCGGACGCGCGGACACATGACCTTGAAGCT	480
Db	576	CTGACGGGCAAGCTGGAACCGGATGCTCCCGGACGCGCGGACACATGACCTTGAAGCT	635
Qy	481	GGAGCTGTGCGCGCACGCGCTGTCCAGAGAGAGCTGATCCAGAACATGAGACCGCGTGA	540
Db	636	GGAGCTGTGCGCGCACGCGCTGTCCAGAGAGAGCTGATCCAGAACATGAGACCGCGTGA	695
Qy	541	CGAGAGATCACCATGTGTAGAGACAGCATCTCTTAAGCTGAAGAGAGACAGACAGCT	600
Db	696	CGAGAGATCACCATGTGTAGAGACAGCATCTCTTAAGCTGAAGAGAGACAGACAGCT	755
Qy	601	GGAGGAGGAGGCTGCGCAACCGCGCGGAGCTGAGAAAGCCCGTGTACCGCGCGCATCGA	660
Db	756	GGAGGAGGAGGCTGCGCAACCGCGCGGAGCTGAGAAAGCCCGTGTACCGCGCGCATCGA	815
Qy	661	GTCCAGAGCACCGCAGCTGTGTCAATCTTACGACGAGAAACCGGAAAGAGCTTGAAGC	720
Db	816	GTCCAGAGCACCGCAGCTGTGTCAATCTTACGACGAGAAACCGGAAAGAGCTTGAAGC	875
Qy	721	TGCACATCGGATTTCTGGAAGGCTGGGGGCCCAAGTGAAGCTGCGCTGTACACAGCC	780
Db	876	TGCACATCGGATTTCTGGAAGGCTGGGGGCCCAAGTGAAGCTGCGCTGTACACAGCC	935

QY 781 CTCGACACCCGGCAGTATCATGAGAACATCAAAATAAAACAGGCGATGCCGAGAGAGCT 840  
Db 996 CTCGACACCCGGCAGTATCATGAGAACTCAAAATAAAACAGGCGATGCCGAGAGAGAGCT 995  
QY 841 AATCTTGTATCTTCAAGAGAGAGAAATCAAGCTGGGAAACAAATGGAGACGAATTTCTGCA 900  
Db 996 AATCTTGTATCTTCAAGAGAGAGAAATCAAGCTGGGAAACAAATGGAGACGAATTTCTGCA 1055  
QY 901 GCGCTATGACCAAGCTCATGAGAGGCTTGGAAAAAAAAGTGGAGCCGATCGAAAAACAACC 960  
Db 1056 GCGCTATGACCAAGCTCATGAGAGGCTTGGAAAAAAAAGTGGAGCCGATCGAAAAACAACC 1115  
QY 961 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGCGAGTACTAGAAAACAGATTCTTGAGAT 1020  
Db 1116 CCGGCGCGCGGCGCAAGAGAGCAAGGTTCGAGTACTAGAAAACAGATTCTTGAGAT 1175  
QY 1021 CCGCAAGCAGCCCGAGCTGACAGAGCGCATGCAAGACAGGTTGGGCGACGCGGCGAGTGG 1080  
Db 1176 CCGCAAGCAGCCCGAGCTGACAGAGCGCATGCGAG--AGGTTGGGCGACGCGGCGAGTGG 1232  
QY 1081 GCTGTTCATGTGCGGCGCGCCGAGCGAGACAGAGGTTCAGAGATCATGATGCGCTCTC 1140  
Db 1233 GCTGTTCATGTGCGGCGCGCCGAGCGAGACAGAGGTTCAGAGATCATGATGCGCTCTC 1292  
QY 1141 AGAGCAGAGAACTTGGAGAGCAAGTCGCGCACTGCGCTGTATCCGCGCATGCTGTA 1200  
Db 1293 AGAGCAGAGAACTTGGAGAGCAAGTGCAGCTGCGCTGTATCCGCGCATGCTGTA 1352  
QY 1201 CGAAGCTGACCAAGCAGCCGATGAAGTTCATCAATGAAACGCGCTTATGGCCGACCCAT 1260  
Db 1353 CGAAGCTGACCAAGCAGCCGATGAAGTTCATCAATGAAACGCGCTTATGGCCGACCCAT 1412  
QY 1261 GAAAGTGTACAAAGACCCGCGAGGTCAATGAAATGTGGAGTGAAGAGAAAGAACTTT 1320  
Db 1413 GAAAGTGTACAAAGACCCGCGAGGTCAATGAAATGTGGAGTGAAGAGAAAGAACTTT 1472  
QY 1321 CCGGAGAAAGTTTCATGACAGCATCCCAAGAACTTTGGCTGATGCAATCTCTGAGAG 1380  
Db 1473 CCGGAGAAAGTTTCATGACAGCATCCCAAGAACTTTGGCTGATGCAATCTCTGAGAG 1532  
QY 1381 GAAAGCAGTGGGTGAGTGGTCTCTATTTACTTACTGACTAAGAAAGAAAGAACTTATA 1440  
Db 1533 GAAAGCAGTGGGTGAGTGGTCTCTATTTACTTACTGACTAAGAAAGAAAGAACTTATA 1592  
QY 1441 GAGCTGTGAGACCGAGCTATTCGCGCGCCGCGCAAGAGCCGACAGCAACAACAGACGA 1500  
Db 1593 GAGCTGTGAGACCGAGCTATTCGCGCGCCGCGCGCAAGAGCCGACAGCAACAACAGACGA 1652  
QY 1501 GCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
Db 1653 GCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1712  
QY 1561 AGATGAGAAAGAGAAAGAAAGAGCGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
Db 1713 AGATGAGAAAGAGAAAGAAAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1772  
QY 1621 CGACAGAGAAAGCTCTCTCAAGAGAAAGCAGACGACCTCAGGAGAGAGACAAACGACGA 1680  
Db 1773 CGACAGAGAAAGCTCTCTCAAGAGAAAGCAGACGACCTCAGGAGAGAGACAAACGACGA 1832  
QY 1681 GAAAGAGCTGTGCTCTCCAAAGGCGCGCAAACTGCGCAACAGCCAGAGAGAGACGCAAGG 1740  
Db 1833 GAAAGAGCTGTGCTCTCCAAAGGCGCGCAAACTGCGCAACAGCCAGAGAGAGACGCAAGG 1892  
QY 1741 CCGCATCACCCGCTCAATGGCTTATAGAGCCCAACAGCGAGAGGCGCATACCCCCCAAGCA 1800  
Db 1893 CCGCATCACCCGCTCAATGGCTTATAGAGCCCAACAGCGAGAGGCGCATACCCCCCAAGCA 1952  
QY 1801 GAGCGCCGAGCTGCGCTCCATGAGAGCTGAATGAGTTCTCGCTGGACAGAGAGAGAAAT 1860  
Db 1953 GAGCGCCGAGCTGCGCTCCATGAGAGCTGAATGAGTTCTCGCTGGACAGAGAGAGAAAT 2012  
QY 1861 GGAACACGCAAGAAAGATCTCTGAGAACAGGCGCAACTGGTGGCCATGCGCCCGAT 1920

Db 2013 GGAACACGCAAGAAAGATCTCTGAGAACAGGCGCAACTGGTGGCCATGCGCCCGAT 2072  
QY 1921 GGTGGGCTCCAAGACTGTGTCGAGTGTAAAGACTTACTTCAACTCAAGAAAGAGCA 1980  
Db 2073 GGTGGGCTCCAAGACTGTGTCGAGTGTAAAGACTTACTTCAACTCAAGAAAGAGCA 2132  
QY 1981 GAACTCGATGATGATCTTTGACAGCAGCAACAGCTGAAGTGAAGAGAGAGAGAGAGCGCG 2040  
Db 2133 GAACTCGATGATGATCTTTGACAGCAGCAACAGCTGAAGTGAAGAGAGAGAGAGAGCGCG 2192  
QY 2041 GAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
Db 2193 GAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2252  
QY 2101 GAGAGATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160  
Db 2253 GAGAGATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2312  
QY 2161 GAGTGAAGCTTTACATGCGCTTGGAGATGAAGTGCAGAGAGAGATGAGTGGCCAGC 2220  
Db 2313 GAGTGAAGCTTTACATGCGCTTGGAGATGAAGTGCAGAGAGAGATGAGTGGCCAGC 2322  
QY 2221 CACTGTCAACAAACAGCTGAGACACGAGAGATCCCTCTCTCAACTGAGAGCGCGCA 2280  
Db 2323 -ACTGTCAACAAACAGCTGAGACACGAGAGATCCCTCTCTCAACTGAGAGCGCGCA 2381  
QY 2281 GACACAGGAGAGAAATGGGCGCAAGCCGCGAGCCCAAGCCCTGAGGCGCGAGCGGCGACCC 2340  
Db 2382 GACACAGGAGAGAAATGGGCGCAAGCCGCGAGCCCAAGCCCTGAGGCGCGAGCGGCGACCC 2441  
QY 2341 AGGCCACCCACCCACCAAGAGAGACATCCGAGCGCCCATGAGCCACCCGCGCTC 2400  
Db 2442 AGGCCACCCACCCACCAAGAGAGACATCCGAGCGCCCATGAGTCAACCCGCGCTC 2501  
QY 2401 TGAAGCCACCGAGCGCTTAAGCCCCCAACAGACCCCATGCGCTTGAACCTGCTCC 2460  
Db 2502 TGAAGCCACCTTAAGCGCTTACGCCCCCAACAGACCCCATTTCTCTTCACTCTCC 2561  
QY 2461 TGTGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
Db 2562 TGTGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2621  
QY 2521 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
Db 2622 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2681  
QY 2581 GCCGCTCAAGAGCTGAGTCAACGAGAGAAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
Db 2682 GCCGCTCAAGAGCTGAGTCAACGAGAGAAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 2741  
QY 2641 GGAAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
Db 2742 GGAAGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2801  
QY 2701 CGGCAAGGCAACCACTGCGCAAGAGCTCGAGGCGCCCGCAGAGACAGCACTCAAGTCTAC 2760  
Db 2802 CGGCAAGGCAACCACTGCGCAAGAGCTCGAGGCGCCCGCAGAGACAGCACTCAAGTCTAC 2861  
QY 2761 CTGCAGTGCAAGCAGAGTGAATGAGGCGGAGGCGGCGCAAGAAACCGGCTGTGTCC 2820  
Db 2862 CTGCAGTGCAAGCAGAGTGAATGAGGCGGAGGCGGCGCAAGAAACCGGCTGTGTCC 2921  
QY 2821 AAGGCCAGGCTCTCAACCCGAGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
Db 2922 AAGGCCAGGCTCTCAACCCGAGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2981  
QY 2881 ACTGAGCTGAAGCAGTGAAGCAGCGAGCGCTGCACTCCCGCATCAAGTCAACCA 2940  
Db 2982 ACTGAGCTGAAGCAGTGAAGCAGCGAGCGGCTGCACTCCCGCATCAAGTCAACCA 3041  
QY 2941 AGTCCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000

Db 3042 AGTCATGACCCCCCGGAGGACGCAAGCTCCACCAAGCACTCCCCAGCCCCACC 3101  
Qy 3001 GCCACGCGAAACCTGACGCGGAGAGCGACGCCCCCTACAGACTTGGACAGCCCCCG 3060  
Db 3102 GCCACGCGAAACCTGACGCGGAGAGCGACGCCCCCTACAGACTTGGACAGCCCCCG 3161  
Qy 3061 GGGCAAGAGCAGGACCGGCAACCCCGCGCAAGAGAG----- 3100  
Db 3162 GGGCAAGAGCAGGACCGGCAACCCCGCGCAAGAGAGCGCTGTGTTCTT 3221  
Qy 3101 ----GCTTTCGACGCGAGGCCCAAGAGCTTCTGGGACCCCCCTTTCGTAATTCGCG 3156  
Db 3222 CCCAGGCTTTCGACGCGAGGCCCAAGAGCTGCTGGGGAGCCCCCTTTCGTAATTCGCG 3281  
Qy 3157 CCGGCTTTCGCGGTCGCGCGCGCGCGCGGAGGATCAAGAGCTTCCCGGATGCGCGAGCC 3216  
Db 3282 CCGGCTTTCGCGGTCGCGCGCGCGCGGAGGATCAAGAGCTTCCCGGATGCGCGAGCC 3341  
Qy 3217 CTGAGCTTTCCTCTAGAGCTCACTGAGTCAACCACTGCGCTTGGGCTTCATGACATGCG 3276  
Db 3342 CTGAGCTTTCCTCTAGAGCTCACTGAGTCAACCACTGCGCTTGGGCTTCATGACATGCG 3401  
Qy 3277 CCGGCTTTCGCGGTCGCGCGCGCGCGCAACCACTTCGCAACCGGCTTCCCTCATTTCTTCGCG 3336  
Db 3402 CCGGCTTTCGCGGTCGCGCGCGCGCGCAACCACTTCGCAACCGGCTTCCCTCATTTCTTCGCG 3461  
Qy 3337 CAAGCAACCCAGAGTCTCGAGAGGCAAAAGGTGCCATCTCCCAAGGAATGTCGAGTCA 3396  
Db 3462 CAAGCAACCCAGAGTCTCGAGAGGCAAAAGGTGCCATCTCCCAAGGAATGTCGAGTCA 3521  
Qy 3397 GCTCCAGTCCCTGACTCAGAGCATGCAAGAGCCCGGCTGGGCTTGTACCACTGGGAGCT 3456  
Db 3522 GCTCCAGTCCCTGACTCAGAGCATGCAAGAGCCCGGCTGGGCTTGTACCACTGGGAGCT 3581  
Qy 3457 GCCCTGCGCCAGTGAACCCCAAAAGCTGGGCACTTCAGCGAGTGAAGCAGGAGCACT 3516  
Db 3582 GCCCTGCGCCAGTGAACCCCAAAAGCTGGGCACTTCAGCGAGTGAAGCAGGAGCACT 3641  
Qy 3517 GTTCCCAACCGGAGGAGGCTTGGGCAACGAGAGCTTGGGAGTCCCAAGCCCAAGAGAGC 3576  
Db 3642 GTTCCCAACCGGAGGAGGCTTGGGCAACGAGAGCTTGGGAGTCCCAAGCCCAAGAGAGC 3701  
Qy 3577 GTTCCGCTGAGAGGAGGAGCTCTGGGCTCAGTTCCGCGGAGAGCACTCAAAAGGCACT 3636  
Db 3702 GTTCCGCTGAGAGGAGGAGCTCTGGGCTCAGTTCCGCGGAGAGCACTCAAAAGGCACT 3761  
Qy 3637 TCCCAAGCAACGAGGTGCTTCGAGCAGCGGCATCACTAACCGGCTCATCAACCAACG 3696  
Db 3762 TCCCAAGCAACGAGGTGCTTCGAGCAGCGGCATCACTAACCGGCTCATCAACCAACG 3821  
Qy 3697 CAGCGCAGCTGAGAGTCTGTGACAAAGGCACTATCAAGAGATCATCGGCGAGAGCAGCCC 3756  
Db 3822 CAGCGCAGCTGAGAGTCTGTGACAAAGGCACTATCAAGAGATCATCGGCGAGAGCAGCCC 3881  
Qy 3757 GAGTCCGCTTGGACCGGCGCGGAGGAGCAGCTGCCCAAGGCGCAAGTCACTTACGAAG 3816  
Db 3882 GAGTCCGCTTGGACCGGCGCGGAGGAGCAGCTGCCCAAGGCGCAAGTCACTTACGAAG 3941  
Qy 3817 CAAGAGAGGCGCAGTCTTGTCTATGAGGAGTGAATGTCTGTGACCAAGTCTTCAAGGA 3876  
Db 3942 CAAGAGAGGCGCAGTCTTGTCTATGAGGAGTGAATGTCTGTGACCAAGTCTTCAAGGA 4001  
Qy 3877 GGAACGGCAGAGAGCTCAGAGACCCCCCATGAGAGGCGCGCGCCCAAGGCGCACTATGA 3936  
Db 4002 GGAACGGCAGAGAGCTCAGAGACCCCCCATGAGAGGCGCGCGCCCAAGGCGCACTATGA 4061  
Qy 3937 CATGATGAGAGGCGCGGAGGAGAGCAGCTCTGAGCAGCAGCATGCAAGGCTCTCATGAG 3996  
Db 4062 CATGATGAGAGGCGCGGAGGAGAGCAGCTCTGAGCAGCAGCATGCAAGGCTCTCATGAG 4121  
Qy 3997 CCGTGCATTCGCGCGGAGGAGCAGAGCCCAACCACTTCAAGAGCAGCAGCACTCG 4056  
Db 4122 CCGTGCATTCGCGCGGAGGAGCAGAGCCCAACCACTTCAAGAGCAGCAGCACTCG 4181

Qy 4057 CCGGTCCATCAACAAGAGATCCCTGCTCTACGTGAGGCAAGAGAGCTACCTGCG 4116  
Db 4182 CCGGTCCATCAACAAGAGATCCCTGCTCTACGTGAGGCAAGAGAGCTACCTGCG 4241  
Qy 4117 TCGGAGGCCAAGCTCTTAAAGCGGAGGAGCAGCTTCGCCCCCAACGCTTCACGGA 4176  
Db 4242 TCGGAGGCCAAGCTCTTAAAGCGGAGGAGGAGCAGCTTCGCCCCCAACGCTTCACGGA 4301  
Qy 4177 CTGACCGAGGCTTACAAAGCAGAGCCCTGAGGCCCCCTGAACTGAAGCGGCGCCATGA 4236  
Db 4302 CTGACCGAGGCTTACAAAGCAGAGCCCTGAGGCCCCCTGAACTGAAGCGGCGCCATGA 4361  
Qy 4237 GGGCTGTGTGCGCAGGAGTGAAGAGGAGGCGGCTCCTCATGAGATCCCGGCGAGGA 4296  
Db 4362 GGGCTGTGTGCGCAGGAGTGAAGAGGAGGCGGCTCCTCATGAGATCCCGGCGAGGA 4421  
Qy 4297 GCTGCGGCAACGCGCGGAGCTGCGCCCGGCGCGGCTCAAGAGGAGCTCATGAC 4356  
Db 4422 GCTGCGGCAACGCGCGGAGCTGCGCCCGGCGCGGCTCAAGAGGAGCTCATGAC 4481  
Qy 4357 GCAAGGCAACCCGCTCAAGTACGACACCGGCGGCTCCACTGCTCCAAAAGCAG 4416  
Db 4482 GCAAGGCAACCCGCTCAAGTACGACACCGGCGGCTCCACTGCTCCAAAAGCAG 4541  
Qy 4417 CGTACGCTCCCTCATGCGGAGCGCCCGGCGGAGCGTTCCCAACCGGTGACCGCGTGAATG 4476  
Db 4542 CGTACGCTCCCTCATGCGGAGCGCCCGGCGGAGCGTTCCCAACCGGTGACCGCGTGAATG 4601  
Qy 4477 GATGCGCAGCGCCCGGCACTGGAACGTGCTGCTACGAGAGAGCTGAAGAGCGGCG 4536  
Db 4602 GATGCGCAGCGCCCGGCACTGGAACGTGCTGCTACGAGAGAGCTGAAGAGCGGCG 4661  
Qy 4537 AGGAGCCGCAAGAGCTCGGAGGAGCTCATTTGCGCGCGGCGCGCTCATTTGCTTGA 4596  
Db 4662 AGGAGCCGCAAGAGCTCGGAGGAGCTCATTTGCGCGCGGCGCGCTCATTTGCTTGA 4721  
Qy 4597 GCTGGGTAAGCGCGGCGGAGCGCCCGTGAACCTTGAAGAACAGGAGGACCTTTGCGCG 4656  
Db 4722 GCTGGGTAAGCGCGGCGGAGCGCCCGTGAACCTTGAAGAACAGGAGGACCTTTGCGCG 4781  
Qy 4657 CCACTTCCACAGAGTTCGCGCTGACATGCGGAGAGCCACGCGCGCTGACAGAGG 4716  
Db 4782 CCACTTCCACAGAGTTCGCGCTGACATGCGGAGAGCCACGCGCGCTGACAGAGG 4841  
Qy 4717 CAGCTTTTGTTCAGAGGAGCATCCAGAGCGGAAGCTGACGTGACGCTTGTGAGAT 4776  
Db 4842 CAGCTTTTGTTCAGAGGAGCATCCAGAGCGGAAGCTGACGTGACGCTTGTGAGAT 4901  
Qy 4777 CCGCAAGTCCCGCAGAGCAGTGCAGAGCAGACCAACCCATCTGCGCTTATGA 4836  
Db 4902 CCGCAAGTCCCGCAGAGCAGTGCAGAGCAGACCAACCCATCTTGTGCTTATGA 4961  
Qy 4837 GCACTGTCTTGGGAGGAGTGAAGTGAAGCTGATTCGAGGCAATCCCTGAGCTT 4896  
Db 4962 GCACTGTCTTGGGAGGAGTGAAGTGAAGCTGATTCGAGGCAATCCCTGAGCTT 5021  
Qy 4897 CGAACCCACTCTCATACCCCGGCGCATTCCTTGAACGAGCGGCTGCTTACTGCGC 4956  
Db 5022 CGAACCCACTCTCATACCCCGGCGCATTCCTTGAACGAGCGGCTGCTTACTGCGC 5081  
Qy 4957 CCGAAGCTTGGGCGCGCAACCCCACTACCGGAGCAGTGAACCCCACTCATTCGCGG 5016  
Db 5082 CCGAAGCTTGGGCGCGCAACCCCACTACCGGAGCAGTGAACCCCACTCATTCGCGG 5141  
Qy 5017 CTACCCCGACAGCGCGCGCTGAGAGAACCGGAGACCATATCATATGACTCATCTTC 5076  
Db 5142 CTACCCCGACAGCGCGCGCTGAGAGAACCGGAGACCATATCATATGACTCATCTTC 5201  
Qy 5077 GCAAGCAGATGACCAACAACGCGCACCGCATGAGCCAGCGAGCTGATATGTGAGGG 5136  
Db 5202 GCAAGCAGATGACCAACAACGCGCACCGCATGAGCCAGCGAGCTGATATGTGAGGG 5261

QY 5137 CCTCTGCCCCGAGTCTCGTGGCACTAAGCTGCGGTCCCCGAGGACATCAT 5196  
Db 5262 CCTCTGCCCCGAGTCTCGTGGCACTCAACTACGCTGCGGGTCCCCGAGGACATCAT 5321  
QY 5137 GCAACCTGCCCCAAGTGGCAACATGCTGCTGCTCGGCCCCCGGACACAGGAGACCCGACG 5256  
Db 5322 GCAACCTGCCCCAAGTGGCAACATGCTGCTGCTCGGCCCCCGGACACAGGAGACCCGACG 5381  
QY 5257 CACCGGCATGAGCCGCTTGCTTACCTCCCAACCGGACCCGACCCCTTGAGAGCGGCA 5316  
Db 5382 CACCGGCATGAGCCGCTTGCTTACCTCCCAACCGGACCCGACCCCTTGAGAGCGGCA 5441  
QY 5317 CAGCAGCTCCCCACTCTCCCAAGAGTCAACAATTGACAAAACAACAACAACGCTC 5376  
Db 5442 CAGCAGCTCCCCACTCTCCCAAGAGTCAACAATTGACAAAACAACAACAACGCTC 5501  
QY 5377 CTGCTCCAGCGGGGACGAGACCGGAGTGGAGAGCGGGACCGGGATCCGGGAGCGGGAAAA 5436  
Db 5502 CTGCTCCAGCGGGGACGAGACCGGAGTGGAGAGCGGGACCGGGATCCGGGAGCGGGAAAA 5561  
QY 5437 GTCCATCTCAAGTCCACCAAGAGTGGAGACGCAACCATCTGGAGACTGTACAGA 5496  
Db 5562 GTCCATCTCAAGTCCACCAAGAGTGGAGACGCAACCATCTGGAGACTGTACAGA 5621  
QY 5497 GCAAGAGCAGCGGACAGACGCGGACAGCGCGGCGGTGGGGGACAGACGCCGCCGCCG 5556  
Db 5622 GCAAGAGCAGCGGACAGACGCGGACAGCGCGGCGGTGGGGGACAGACGCCGCCGCCG 5681  
QY 5557 CTCCCACTCCCAATGCCACCAACTCGGCCCATCTCCCTCCGACCCAGAGAGCCCTCA 5616  
Db 5682 CTCCCACTCCCAATGCCACCAACTCGGCCCATCTCCCTCCGACCCAGAGAGCCCTCA 5741  
QY 5617 GCAGAGACCAGTGTGCTTCAACAACAGAGCATGAAGGTATCATCACCGCTGGAGGC 5676  
Db 5742 GCAGAGACCAGTGTGCTTCAACAACAGAGCATGAAGGTATCATCACCGCTGGAGGC 5801  
QY 5677 CAGCAAGCCCAAGTCTTGAAGTTCACCTTCACCTTCACCCGTTGCGCCAGTGGCAC 5736  
Db 5802 CAGCAAGCCCAAGTCTTGAAGTTCACCTTCACCTTCACCCGTTGCGCCAGTGGCAC 5861  
QY 5737 ATTCCCACTGCGACCACTGCGCACTGGGCGGCAACCTCGATGGGGTTCACCTTACCT 5796  
Db 5862 ATTCCCACTGCGACCACTGCGCACTGGGCGGCAACCTCGATGGGGTTCACCTTACCT 5921  
QY 5797 CATGAGCCCGCTTGTGCTGCCCCAAGAGGCCCCCGGAGTGGCCGAGCAGAGCGGCCG 5856  
Db 5922 CATGAGCCCGCTTGTGCTGCCCCAAGAGGCCCCCGGAGTGGCCGAGCAGAGCGGCCG 5981  
QY 5857 AGCAGACACCGGCAATGCTTCTTCGCCAAGCCCCCAAGCCCGCTCGGAGTGGAGCCGC 5916  
Db 5982 AGCAGACACCGGCAATGCTTCTTCGCCAAGCCCCCAAGCCCGCTCGGAGTGGAGCCGC 6041  
QY 5917 CTGCTCCCGCAGAGAGGCTGCGAGCCCCGAGCCCTTAAAGTCTCTGTCTTGGCAGCG 5976  
Db 6042 CTGCTCCCGCAGAGAGGCTGCGAGCCCCGAGCCCTTAAAGTCTCTGTCTTGGCAGCG 6101  
QY 5977 CACCACTGCGGCAACCCCTGCGAAGAACTCGCACTTCACAGCAGCGCCGAGCCCGCC 6036  
Db 6102 CACCACTGCGGCAACCCCTGCGAAGAACTTCGCACTTCACAGCAGCGCCGAGCCCGCC 6161  
QY 6037 GAGCGCACTGCTGCTGAGCTCGAGCCCGGACCGGAGAAAGTCAAAAGTAAACCTTTTTC 6096  
Db 6162 GAGCGCACTGCTGCTGAGCTCGAGCCCGGACCGGAGAAAGTCAAAAGTAAACCTTTTTC 6221  
QY 6097 CATTCAGAGAACTGGAACTCCGTTCTTGGGTTAACACAGGCAAGATTAAGCCCCGAGG 6156  
Db 6222 CATTCAGAGAACTGGAACTCCGTTCTTGGGTTAACACAGGCAAGATTAAGCCCCGAGG 6281  
QY 6157 GGTGAGGCCCGTCAAGCCCTGTAGAGTCAACCAAGTTCAGCCAGCAAGAGGGCTCCCAA 6216  
Db 6282 GGTGAGGCCCGTCAAGCCCTGTAGAGTCAACCAAGTTCAGCCAGCAAGAGGGGCTCCCAA 6341  
QY 6217 GCACTTGAAGAGTTCAGCAAGAGCCACTTGAAGGGGAGCTTGGCGCCCAAGAGCCAGG 6276

Db 6342 GCACTTGAAGAGTTCAGCAAGAGCCACTTGAAGGGGAGTCTGCGGCCCAAGCAACGAG 6401  
QY 6277 CCCCCTGAAGCTTGGCGGGGAGCGGCCCACTTCCACACACTTGCAGGCGCTGAGAG 6336  
Db 6402 CCCCCTGAAGCTTGGCGGGGAGCGGCCCACTTCCACACACTTGCAGGCGCTGAGAG 6461  
QY 6337 CCAAGCCTTGTTCAGCCCGCTGCTCAGACCGGCCCAAGGGGTCAAAAGTCAACAGCGGT 6396  
Db 6462 CCAAGCCTTGTTCAGCCCGCTGCTCAGACCGGCCCAAGGGGTCAAAAGTCAACAGCGGT 6521  
QY 6397 GGTCACTTGGGCTCCAGCACAATGAGAGTATCAACAGAGCTTACCCCGACACACC 6456  
Db 6522 GGTCACTTGGGCTCCAGCACAATGAGAGTATCAACAGAGCTTACCCCGACACACC 6581  
QY 6457 ACAGAGCTCAGCGACACCCCTGCGCCCGCCCTTACCTCTTCCCTGGGGGACAGCTGCC 6516  
Db 6582 ACAGAGCTCAGCGACACCCCTGCGCCCGCCCTTACCTCTTCCCTGGGGGACAGCTGCC 6641  
QY 6517 GGTCTTGAACCTTCGCGCGCCACCCAGTGAACCTTACCTCCGCCCCGAGCAATGTGC 6576  
Db 6642 GGTCTTGAACCTTCGCGCGCCACCCAGTGAACCTTACCTCCGCCCCGAGCAATGTGC 6701  
QY 6577 CCGGCGCTGCTGCTCCCCCAGAGGAGGGGCAAGAGTTCAGAGAGCCAAACAAGAC 6636  
Db 6702 CCGGCGCTGCTGCTCCCCCAGAGGAGGGGCAAGAGTTCAGAGAGCCAAACAAGAC 6761  
QY 6637 GTGCTTGGGTGGTGGAGACGATTTGAACCTGTCCTCCACCGAGAGGATGAC 6696  
Db 6762 GTGCTTGGGTGGTGGAGACGATTTGAACCTGTCCTCCACCGAGAGGATGAC 6821  
QY 6697 GAGGCCAGGGCACTCCGAGAGTGTGTGACCCGCTGCTGACCGGAGTGGGAAACAAC 6756  
Db 6822 GAGGCCAGGGCACTCCGAGAGTGTGTGACCCGCTGCTGACCGGAGTGGGAAACAAC 6881  
QY 6757 GAGGCCAGAGAGTGGGCTCCAAATCTTCAGGCAACACAGCCAGCGGCCAGCTTCTT 6816  
Db 6882 GAGGCCAGAGAGTGGGCTCCAAATCTTCAGGCAACACAGCCAGCGGCCAGCTTCTT 6941  
QY 6817 CAGCAAGCTGACCCGAGACACTCCGCATGGTCAAGTCCAAAGAACAAAGATCAACA 6876  
Db 6942 CAGCAAGCTGACCCGAGACACTCCGCATGGTCAAGTCCAAAGAACAAAGATCAACA 7001  
QY 6877 GAAAGTGAACAACCAACCGGAAATGAGCTGAATATACATATACAGCCAGCTGGAGCGA 6936  
Db 7002 GAAAGTGAACAACCAACCGGAAATGAGCTGAATATACATATACAGCCAGCTGGAGCGA 7061  
QY 6937 GATCTTCAATATGCCCCGATACACGGAAACAGGCTTATGACTTATAGAGGCAAGCGGT 6996  
Db 7062 GATCTTCAATATGCCCCGATACACGGAAACAGGCTTATGACTTATAGAGGCAAGCGGT 7121  
QY 6997 GCAGGAATGTCAGACACAACATGGGGCTGAGGCGCAATATTAGAAAGGCACTATGGG 7056  
Db 7122 GCAGGAATGTCAGACACAACATGGGGCTGAGGCGCAATATTAGAAAGGCACTATG-- 7179  
QY 7057 TAAATATGACAGTGGGAAAGTCCCGCGCTCAGCGCAATGTCTTTAAACCTCTGAA 7116  
Db 7180 ----- 7179  
QY 7117 TGCCAGTGCAGCTGCGCGCTGTATGCCCATTAACCGCTGTGACGAGCGAGTGACCA 7176  
Db 7180 ----- 7179  
QY 7177 CACACTCACTTCGCAAGGTGGGAGGAGGAGCCAAAGGTCTTGGCAGACCCAGAGCGG 7236  
Db 7180 -----GATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7223  
QY 7237 AAAAGCAAGTCCCGGAGCCCGAGGCTGGCAATCTGGGAGACCGGACCACTGTCTCTC 7296  
Db 7224 AAAAGCAAGTCCCGGAGCCCGAGGCTGGCAATCTGGGAGACCGGACCACTGTCTCTC 7283  
QY 7297 AGTGCATTCGAGGAGGAGCTGCAACCGCGGACCGCGCTCAACAACCGGCTGTGAGAGG 7356



Dh	7284	AGTGCATCTGGAGGGGAGACTGCAACCGCCGGAGCGCCCTCAACCAACCGCGTGTGGAGGA	7343
Oy	7357	CAGGCGCTGTCTCCGAGGGTTCACGCGCATTCCTCTTACAACCCCTCGATCATGCGGCTGCA	7416
Dh	7344	CAGGCGCTGTCTCCGAGGGTTCACGCGCATTCCTCTTACAACCCCTCGATCATGCGGCTGCA	7403
Oy	7417	GGCGGGGTGTCAATGGCTTTCCTCCACCCCGGAGCTCCCCGGGGCAAGCGGGCCCTTCGCG	7476
Dh	7404	GGCGGGGTGTCAATGGCTTTCCTCCACCCCGGAGCTCCCCGGGGCAAGCGGGCCCTTCGCG	7463
Oy	7477	TGGCCCCCACAACCGCTGGGAGCGAGGAGCCCAAGGACCTGCGTCTGTGCGAGTACAGAGAC	7536
Dh	7464	TGGCCCCCACAACCGCTGGGAGCGAGGAGCCCAAGGACCTGCGTCTGTGCGAGTACAGAGAC	7523
Oy	7537	ACTCTCCGACAGCGAGTGACTCAGAACAGGGCGGGGGGGGGCGGGCGGTBUERSPERLUG	7596
Dh	7524	ACTCTCCGACAGCGAGTGACTCAGAACAGGGCGGGGGGGGGCGGGCGGTBUERSPERLUG	7573
Oy	7597	TCAGGTCCCAAGCGACCAAGAGAACGGCCCTGCAAGAGCGGGGGGGGGCTGCGGACTCCCC	7656
Dh	7574	TCAGGTCCCAAGCGACCAAGAGAACGGCCCTGCAAGAGCGGGGGGGGGCTGCGGACTCCCC	7633
Oy	7657	AACCAAGGAAGAAGCGCCCTGAGTCCGCGCTGCGGCGCTCATTCATCTGTCCGTCAGAGCGCG	7716
Dh	7634	AACCAAGGAAGAAGCGCCCTGAGTCCGCGCTGCGGCGCTCATTCATCTGTCCGTCAGAGCGCG	7693
Oy	7717	GCAATCTTTCGCTGTCTTAAAGCTTTAACTTAAGCTCCCGCCCGGGGCTTGCCCTGTGCGAGA	7776
Dh	7694	GCAATCTTTCGCTGTGTCTTAAAGCTTTAACTTAAGCTCCCGCCCGGGGCTGTGTGCGAGA	7753
Oy	7777	CCTTAATCAGGGGAGATGTTTACCTGGTGTCTGGGAAAGGAGGGGAAAGGGGCGGAGAGGGG	7836
Dh	7754	CCTTAATCAGGGGAGATGTTTACCTGGTGTCTGGGAAAGGAGGGGAAAGGGGCGGAGAGGGG	7813
Oy	7837	GCAAGGGAGGGGCTGTGTGGAGGCAACACAAAGCGCGCAAGGGGGCGCAGGGAAACCAAGACAG	7896
Dh	7814	GCAAGGGAGGGGCTGTGTGGAGGCAACACAAAGCGCGCAAGGGGGCGCAGGGAAACCAAGACAG	7873
Oy	7897	GATGACACAGCACTTCCAGGCACTGCGCTCCCGCGAATGATCTTGGAAACCAAGACTTTAAA	7956
Dh	7874	GATGACACAGCACTTCCAGGCACTGCGCTCCCGCGAATGATCTTGGAAACCAAGACTTTAAA	7933
Oy	7957	CTGAGCTTCAGAGCCCGCCGCGGCTTCCTCCGCTCCCATCCCGCTTAAAGCTGTGAGCAG	8016
Dh	7934	CTGAGCTTCAGAGCCCGCCGCGGCTTCCTCCGCTCCCATCCCGCTTAAAGCTGTGAGCAG	7993
Oy	8017	ATGGAACGAGGCGCTGTGTCAACCCCCCAAGTGGGCTGTTCGCGTCCCGACAGACTGCCCCA	8076
Dh	7994	ATGGAACGAGGCGCTGTGTCAACCCCCCAAGTGGGCTGTTCGCGTCCCGACAGACTGCCCCA	8053
Oy	8077	GCCAAACGAGATTGCTGCGGAAACCAAGTCAAGGCGCAGGTGGGCGGACAAAGAGGCGCAGGTGCG	8136
Dh	8054	GCCAAACGAGATTGCTGCGGAAACCAAGTCAAGGCGCAGGTGGGCGGACAAAGAGGCGCAGGTGCG	8113
Oy	8137	GCCGTGGGGGAAACGAGATGCTCCGAGGACTGTGACTGTTTTTTTCAACAATCGTTGCCGAG	8196
Dh	8114	GCCGTGGGGGAAACGAGATGCTCCGAGGACTGTGACTGTTTTTTTCAACAATCGTTGCCGAG	8173
Oy	8197	CGGTGGGAAAGAAAGGCGAGATGTAAATGTGTGTGGTTTACAGGGTAATATTTTGTATAC	8256
Dh	8174	CGGTGGGAAAGAAAGGCGAGATGTAAATGTGTGTGGTTTACAGGGTAATATTTTGTATAC	8233
Oy	8257	CTTCAATGAATTAATTAATTCAGATGTTTTTACGAAAGGAAGGACTTACCCAGTATTAATCGCTGC	8316
Dh	8234	CTTCAATGAATTAATTAATTCAGATGTTTTTACGAAAGGAAGGACTTACCCAGTATTAATCGCTGC	8293
Oy	8317	TGTGCTTTTGATCTCTGCTTACCGTTCAAGAGGCGTGTGCAAGGCGCGACAGTGTGATACCC	8376
Dh	8294	TGTGCTTTTGATCTCTGCTTACCGTTCAAGAGGCGGTGTGCAAGGCGCGACAGTGTGATACCC	8353
Oy	8377	CATCATCTGCGAGAACAAAGGGGGGGGGGACTGTCTGTACAGCCCGCGTGTGTCTCTCCTC	8436
Dh	8354	CATCATCTGCGAGAACAAAGGGGGGGGGGACTGTCTGTACAGCCCGCGTGTGTCTCTCCTC	8413

QY	8437	CTCCCTTCTTGGGCGAATGAATTCGATTCGTGGCCCATTTGGCGCGG	8496
Db	8414	CTCCCTTCTTGGGCGAATGAATTCGATTCGTGGCCCATTTGGCGCGG	8473
QY	8497	TGGGTGATTCGTGATTTACACAGCGTCTTAATTAAGGAATTAATCTCAAA	8556
Db	8474	TGGGTGATTCGTGATTTACACAGCGTCTTAATTAAGGAATTAATCTCAAA	8553
QY	8557	AAAAAAAAAAAAAA 8571	
Db	8534	AAAAAAAAAAAAAA 8548	
RESULT 5	HSU37146	5989 bp	mRNA
LOCUS	HSU37146		linear
DEFINITION	Human silencing mediator of retinoid and thyroid hormone action		
ACCESSION	U37146		
VERSION	U37146.1		
KEYWORDS	GI:1045654		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 5989)		
TITLE	Chen, J.D. and Evans, R.M.		
JOURNAL	A transcriptional co-repressor that interacts with nuclear hormone receptors		
PUBMED	Nature 377 (6548), 454-457 (1995)		
REFERENCE	2 (bases 1 to 5989)		
AUTHORS	Chen, J.D. and Evans, R.M.		
JOURNAL	Direct Submission		
FEATURES	Submitted (27-SEP-1995) J. Don Chen, Gene Expression Lab, The Salk Institute, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA		
source	location/Qualifiers		
	1..5989		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/cell_line="Hela"		
	1..5989		
	/gene="SMRT"		
	496..4983		
	/gene="SMRT"		
	/note="transcriptional co-repressor"		
	/codon_start=1		
	/product="silencing mediator of retinoid and thyroid hormone action"		
	protein_id="AAC50226.1"		
	/db_xref="GI:1045655"		
	/translation="MEAMDMPPDKAFAEAQQLPGDPFCWTSGLPFPVPREVIKAS PHADBSAFSAPBPHPLPLGLHDTPARPVLPRTTISNPPLISSAGHPSVLROIGA ISGMSYQLHVPYSEBAKAVGVWTGLPLMPDPKCLAPSGVKOEQLSRGAGPPE SLCPVPAQBAASVLRGALSGPGLSTIKGIISRVSDDAITYRGSTTHETPDVLYK GTTTRIIIGDSESRLDGRDLSLPKHVITYEKKGHVLSLEGMSVYQCKEGRSS GPHETVAEPRTYDMMEGRVGRAISSASIEGLMGRAIPERHSPIHLKEOHNRGSIT QGIIPRSVIAOEADYLRREALRKREBPPPPSRDITAYKYQALGPLKKAHEGL VAATYKAGRSIHIEPREELHTPELPLAPRLKEGSLTQCTPLKYDTGASTGSKID VRSILSGRTPPVPVLDVMDADARALERYCESLSKRSRGTSAGSAGSARGAIVYV PELGRKPSQPIVYEDHGAQCHLPGSGPVYMBEPRLOEGSLSSKSDKLTST PREIASSPHSTVPEHHHPISPIEHLIRGVSGVDLIRSHIPLAFDPISIRGIRPLDAA AAATYPLHLAHPNTPYPLHYPPYLIRGIPDTALENQTIINDYTSQOMHNATYAA QORDMSIRGLSPRESSLALNVAAGPRGIDLSQVPHLPVLPPTPGATMDRLAYLP TAQPDFSRSSSPLSPGCTYHLTKPTTSSSRERDRERDRERERAKSLTSTYT VEHAPIMRPEETEOSSGSSGSSSRPASHSHOHSPISPRDOLLOORPSVL HNGCMGITAVAPSKPVLRSTSSPVAPRATTPDATHCPIGTLIDGVYPTIMPEV LLEKAPRVARPERPRADTGHAFLAPPAASGIEPASPSSPKSGEPRPLVPVVGATII LIRPAKNALHAASPDPAAPASDPHEKTSKPSIDELRLSGITGSSVSPG VEAPSPVSSSLTHDKGLPKHLBELKSLHGLRPQDPPVVLGSGEALHLPLRLP		



ESQSSSPLLQTAPEVKHQVVTLAQHSIVITODYTRHHPQULSAPLPADLYSPFG  
ASCPVLDIRRPSDYLPPEPDHGAPEKSPHEGKRSPEPKTIVLGGEGIEPVS  
PPEGMTEBHSRAVYPLPYRDGEQTEPBMGSKSPGMSQPAFESKLTENSAMVX  
SKKQIKKNTKHNHNEBPYNI SOPETL FNNPALTGMLTRYRQOA VOEHSTJMG  
EALIRKALMGKXDOMESPPLSANA PNPMLNASI.PAMPPLTAAAGSDHLLTSPGG  
GKAKTSGRSPSSKAKSPAPGLASGRPPSVSSVHSEGDENRTPLITNMYMERPSAG  
STPFPYNPILMRLQGMVMSPPPGLPAGSGPLAGHHAMDEPKPLCSQYETLSDS  
E"

4483..4620  
/gene="SMRT"  
/note="alternatively spliced insert"

exon  
Query Match 63.7%; Score 5455; DB 8; Length 5989;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 5477; Conservative 1; Mismatches 5; Indels 10; Gaps 1;

QY	3079	GGCACCCCCCGCCGACAGAGAGGCTTGGCAGCCGAGGCCGAGAAAGTCCCTGGGAGACCC	3138
Db	507	GGACGCCCAACCCGACAGAGAGGCTTGGCAGCCGAGGCCGAGAAAGTCCCTGGGAGACCC	566
QY	3139	CCCTTGCTGAGACTTCGCGGCTGCTCCCTTCCCGGACCCCGGTGAGGTGATCAAGGCTCTC	3198
Db	567	CCCTTGCTGAGACTTCGCGGCTGCTCCCTTCCCGGACCCCGGTGAGGTGATCAAGGCTCTC	626
QY	3199	CCCGATGCCCCCGGACCCCTTCAAGCTTCTCTAAGCTTCAAGCTTCAAGCTTCAAGCTT	3258
Db	627	CCCGATGCCCCCGGACCCCTTCAAGCTTCTCTAAGCTTCAAGCTTCAAGCTTCAAGCTT	686
QY	3259	GGGCTTCCATGACTGCCCCGCGGCTCTGCGCGGCGCCCAACCACTATCCAAAGCCGCTC	3318
Db	687	GGGCTTCCATGACTGCCCCGCGGCTCTGCGCGGCGCCCAACCACTATCCAAAGCCGCTC	746
QY	3319	TCCCTCATCTCTCTGCGCAAGCACCCGAGCGTCTCGAGAGGCAAAATAGTGCCATCTC	3378
Db	747	TCCCTCATCTCTCTGCGCAAGCACCCGAGCGTCTCGAGAGGCAAAATAGTGCCATCTC	806
QY	3379	CCAGGAATGTGGTCCAGCTTCAGGTCCGTACTCAAGACATGCCAAAGCCCGGTGGG	3438
Db	807	CCAGGAATGTGGTCCAGCTTCAGGTCCGTACTCAAGACATGCCAAAGCCCGGTGGG	866
QY	3439	CCCTGTCAACATGGGGCTGCCCCCTGCCATGAGACCCCAAAAAGCTTGACCCCTTCAAGCG	3498
Db	867	CCCTGTCAACATGGGGCTGCCCCCTGCCATGAGACCCCAAAAAGCTTGACCCCTTCAAGCG	926
QY	3499	AGTGAAGCAGAGAGCTGTCCCAACGCGGCGCAGGCTGGGCGCACCGAGAGGCTGGGGT	3558
Db	927	AGTGAAGCAGAGAGCTGTCCCAACGCGGCGCAGGCTGGGCGCACCGAGAGGCTGGGGT	986
QY	3559	GGCCACAGCCGAGAGCGTCCGTGCTGAGAGGAGACAGCTCTGGGCTCAGTTCCGGGCGG	3618
Db	987	GGCCACAGCCGAGAGCGTCCGTGCTGAGAGGAGACAGCTCTGGGCTCAGTTCCGGGCGG	1046
QY	3619	AAGCATCAACAAAGGATTCCTCAAGACAGGGTGCTTGGACAGGCGCATCATACCG	3678
Db	1047	AAGCATCAACAAAGGATTCCTCAAGACAGGGTGCTTGGACAGGCGCATCATACCG	1106
QY	3679	CGGCTTCATCAACCGAGGACCGCAGCTGACGCTCTGTAACAAGGACCATCAAGAGAT	3738
Db	1107	CGGCTTCATCAACCGAGGACCGCAGCTGACGCTCTGTAACAAGGACCATCAAGAGAT	1166
QY	3739	CATCGCGAGAGACAGCCCGAGTCCGCTTGGACCGCGGCGCGGAGAGACAGCTGCCAAAGG	3798
Db	1167	CATCGCGAGAGACAGCCCGAGTCCGCTTGGACCGCGGCGCGGAGAGACAGCTGCCAAAGG	1226
QY	3799	CCAGCTCATCTACGAAGGCAAGAGGCGCAAGCTTTGTCTTATGAGGGTGGCATGTCTGT	3858
Db	1227	CCAGCTCATCTACGAAGGCAAGAGGCGCAAGCTTTGTCTTATGAGGGTGGCATGTCTGT	1286
QY	3859	GACCCAGTGTCCAGAGAGAGGCGGACAGACGCTCAAGGACCCCGCATGAGAGCGGCGCG	3918
Db	1287	GACCCAGTGTCCAGAGAGAGGCGGACAGACGCTCAAGGACCCCGCATGAGAGCGGCGCG	1346

QY	3919	CCCAAGCGCACCTTATGACATGATGAGAGGCGCGCTGGCGAGAGCCATCTCTCAAGCCAG	3978
Db	1347	CCCAAGCGCACCTTATGACATGATGAGAGGCGCGCTGGCGAGAGCCATCTCTCTCAAGCCAG	1406
QY	3979	CATGAAGTCTCAAGGCGCGGACATCCCGCGGAGAGGACACAGCCCGCACCATCTCA	4038
Db	1407	CATGAAGTCTCAAGGCGCGGACATCCCGCGGAGAGGACACAGCCCGCACCATCTCA	1466
QY	4039	AGAGGAGACACACATCCGCGGCTCCATGACACAGAGATCCCTCGGCTCTTACGTGAGGC	4098
Db	1467	AGAGGAGACACACATCCGCGGCTCCATGACACAGAGATCCCTCGGCTCTTACGTGAGGC	1526
QY	4099	ACAGAGAGACTACCTGCGTCCGAGAGCCAGTCTCTTAAAGCGGAGAGGACGCTCCGCTC	4158
Db	1527	ACAGAGAGACTACCTGCGTCCGAGAGCCAGTCTCTTAAAGCGGAGAGGACGCTCCGCTC	1586
QY	4159	CCACCGGCTTCAAGGAGCTGACCGAGGCTCAAGAGCGAGGCGCTGGGCGCCCTGAA	4218
Db	1587	CCACCGGCTTCAAGGAGCTGACCGAGGCTCAAGAGCGAGGCGCTGGGCGCCCTGAA	1646
QY	4219	GCTGAAGCGGCGCATGAGGCGCTGTGCGCACGCTGAAAGAGAGCGGCGCTCCATCA	4278
Db	1647	GCTGAAGCGGCGCATGAGGCGCTGTGCGCACGCTGAAAGAGAGCGGCGCTCCATCA	1706
QY	4279	TGAGATCCCGCGCGAGAGCTGCGGACACGCGCGAGCTGCGCTTGGCTCCGCGCGCT	4338
Db	1707	TGAGATCCCGCGCGAGAGCTGCGGACACGCGCGAGCTGCGCTTGGCTCCGCGCGCT	1766
QY	4339	CAAGAGGCTCATCAAGGAGGAGCCCGCTCAAGTACAGACACCGGCGCTCCACAC	4398
Db	1767	CAAGAGGCTCATCAAGGAGGAGCCCGCTCAAGTACAGACACCGGCGCTCCACAC	1826
QY	4399	TGGCTCAAAAAGCAGAGTACGCTCCCTCATTCGCGAGCCCGGCGGACGTTCCAC	4458
Db	1827	TGGCTCAAAAAGCAGAGTACGCTCCCTCATTCGCGAGCCCGGCGGACGTTCCAC	1886
QY	4459	GCTGACCCCGCTGATGTGATGCGGACGCGCGGACCTGGAAGTCTGCTTACAGAGA	4518
Db	1887	GCTGACCCCGCTGATGTGATGCGGACGCGCGGACCTGGAAGTCTGCTTACAGAGA	1946
QY	4519	GAGCCTGAAGAGCGGCGCAGAGGACCGGACGAGCTCGGGGGCTCATTTGCGGCGGCG	4578
Db	1947	GAGCCTGAAGAGCGGCGCAGAGGACCGGACGAGCTCGGGGGCTCATTTGCGGCGGCG	2006
QY	4579	CCCGTCAATTGTGCTGAGTGTGATAAGCCGCGGACAGACCCCTTGAACCTATAGAGACA	4638
Db	2007	CCCGTCAATTGTGCTGAGTGTGATAAGCCGCGGACAGACCCCTTGAACCTATAGAGACA	2066
QY	4639	CGGGACCCCTTGTGCGGACCTTCCACAGAGTTGCGCGTGAACATGCGGAGGCCAC	4698
Db	2067	CGGGACCCCTTGTGCGGACCTTCCACAGAGTTGCGCGTGAACATGCGGAGGCCAC	2126
QY	4699	GGCGGCTGACAGAGGCGAGCTTTCGTCCAGAGGATCCCGAGACCGGAAAGCTGAC	4758
Db	2127	GGCGGCTGACAGAGGCGAGCTTTCGTCCAGAGGATCCCGAGACCGGAAAGCTGAC	2186
QY	4759	GTGAGCGCTGATGATGCGCAAGTCCCGGACAGACACCGTGGCCGAGACACACCA	4818
Db	2187	GTGAGCGCTGATGATGCGCAAGTCCCGGACAGACACCGTGGCCGAGACACACCA	2246
QY	4819	CCCATCTTGCTTATGAGACCTTCTTGAGGCGGTGAGTGGCTGTATATGCGAG	4878
Db	2247	CCCATCTTGCTTATGAGACCTTCTTGAGGCGGTGAGTGGCTGTATATGCGAG	2306
QY	4879	CCAGATCCCGCTGAGCTTGAACCCCACTCAACCCCGGCGCATCCCTTGAAGGCGAGC	4938
Db	2307	CCAGATCCCGCTGAGCTTGAACCCCACTCAACCCCGGCGCATCCCTTGAAGGCGAGC	2366
QY	4939	CGCTGCTACTACCTGCGCCGACACTTGGGCCCAACCCCACTACCGGCACTGTATCC	4998
Db	2367	CGCTGCTACTACCTGCGCCGACACTTGGGCCCAACCCCACTACCGGCACTGTATCC	2426
QY	4999	ACCTTACTCATCCGCGGCTACCCGAGACGCGCGGCGCTTGAAGAACCGGACAGACATCAT	5058

|||||  
Db 2427 ACCCTACTATCGCGCGGATACCCGACAGCGCGCGCTGAGAAACGGGACACATCAT 2486  
QY 5059 CAATGACTATACATCACTCCGAGAGATGACACAACAACGGCCACCCGCAATGGCCACG 5118  
Db 2487 CAATGACTATACATCACTCCGAGAGATGACACAACAACGGCCACCCGCAATGGCCACG 2546  
QY 5119 AGCTGATATGCTGAGGGGCTCTCGCCCGAGTCTCGCTGGCACTCAACTACGTGC 5178  
Db 2547 AGCTGATATGCTGAGGGGCTCTCGCCCGAGTCTCGCTGGCACTCAACTACGTGC 2606  
QY 5179 GGGTCCCCGAGGATCATGCACTGTCTCCAAATGACACACTGCTGTGTCTGTGCC 5238  
Db 2607 GGGTCCCCGAGGATCATGCACTGTCTCCAAATGACACACTGCTGTGTCTGTGCC 2666  
QY 5239 GACACACAGGACCCCGACCCGACATGACACCGCTTGTCTACTCTCCGACCGGCCCA 5298  
Db 2667 GACACACAGGACCCCGACCCGACATGACACCGCTTGTCTACTCTCCGACCGGCCCA 2726  
QY 5299 GCCCTTCAGAGCGGACACAGACGCTCCCACTCTCCGAGAGTCAACACACTTGAC 5358  
Db 2727 GCCCTTCAGAGCGGACACAGACGCTCCCACTCTCCGAGAGTCAACACACTTGAC 2786  
QY 5359 AAAACCAACCAACGCTCTGTCTCGAGCGGAGACCGGATTCAGAGCGGACCG 5418  
Db 2787 AAAACCAACCAACGCTCTGTCTCGAGCGGAGACCGGATTCAGAGCGGACCG 2846  
QY 5419 GGAATCGGAGCGGAAAAAGTCTCTCAAGTCAACACAGCGGTGAGACACCAT 5478  
Db 2847 GGAATCGGAGCGGAAAAAGTCTCTCAAGTCAACACAGCGGTGAGACACCAT 2906  
QY 5479 CTGAGACCTGTATCAGAGACGAGCGGACAGAGCGGACGCGGAGGAGG 5538  
Db 2907 CTGAGACCTGTATCAGAGACGAGCGGACAGAGCGGACGCGGAGGAGG 2966  
QY 5539 CAGCAGACGCGCGCGCTCTCCCACTCTCCATGCCACAGACACTGCGCATCTCCCTCG 5598  
Db 2967 CAGCAGACGCGCGCGCTCTCCCACTCTCCATGCCACAGACACTGCGCATCTCCCTCG 3026  
QY 5599 GACCCAGATGCGCTCTCAGAGACACCAAGTGTCTTCAACAACAAGGATGAAAGGTAT 5658  
Db 3027 GACCCAGATGCGCTCTCAGAGACACCAAGTGTCTTCAACAACAAGGATGAAAGGTAT 3086  
QY 5659 CATCACGCGTGTGAGGCGGACGAGCGGACGAGTCTGAGGTCCACTGCACTCTCTACC 5718  
Db 3087 CATCACGCGTGTGAGGCGGACGAGCGGACGAGTCTGAGGTCCACTGCACTCTCTACC 3146  
QY 5719 CGTTGCGGACGTCGACATTTCCACTCTGCAACCACTGCGCACTGCGGCGGCACTCTCGA 5778  
Db 3147 CGTTGCGGACGTCGACATTTCCACTCTGCAACCACTGCGCACTGCGGCGGCACTCTCGA 3206  
QY 5779 TGGGGTCTACCTTACCTCTCAATGAGCGGCTGTGTGCGCAAGAGGCGCGCGGCTGCG 5838  
Db 3207 TGGGGTCTACCTTACCTCTCAATGAGCGGCTGTGTGCGCAAGAGGCGCGCGGCTGCG 3266  
QY 5839 CGGCGCAGAGCGGCGCGGACGAGCAACCGGCAATGCTTCTCTGCGCAAGGCGCGGCGG 5898  
Db 3267 CGGCGCAGAGCGGCGCGGACGAGCAACCGGCAATGCTTCTCTGCGCAAGGCGCGGCGG 3326  
QY 5899 CTTCGGGCTGAGACCGGCTCTCTCCCGACAGAGGCTCGAGACCGCGGCGCTTGTGC 5958  
Db 3327 CTTCGGGCTGAGACCGGCTCTCTCCCGACAGAGGCTCGAGACCGCGGCGCTTGTGC 3386  
QY 5959 TCTGTCTCTGCGACGACCACTGCGCGGACCTCTGCGAAGAACTCGCACTTCAACA 6018  
Db 3387 TCTGTCTCTGCGACGACCACTGCGCGGACCTCTGCGAAGAACTCGCACTTCAACA 3446  
QY 6019 CGCAGACCGGACCGCGCGGCGCACCTGCTGCGCTGAGACCGCGACCGGAAAAAGAC 6078  
Db 3447 CGCAGACCGGACCGCGCGGCGCACCTGCTGCGCTGAGACCGCGACCGGAAAAAGAC 3506  
QY 6079 TCAAGTAAACCTTTTCCATCCAGAACTGAACTCCGTTCTCTGGGTTACACGCGCAG 6138

Db 3507 TCAAGTAAACCTTTTCCATCCAGAACTGAACTCCGTTCTCTGGGTTACACGCGCAG 3566  
QY 6139 CAGCTACAGCCCGGAGAGGGGTGAGGCCGCTCAGCGCTGTAGACTCACAGTGTGACCA 6198  
Db 3567 CAGCTACAGCCCGGAGAGGGGTGAGGCCGCTCAGCGCTGTAGACTCACAGTGTGACCA 3626  
QY 6199 CGACAGGGGCTCCCAAGCACTGGAAGACTGACAGAGGCACTTGAAGGGGAGCT 6258  
Db 3627 CGACAGGGGCTCCCAAGCACTGGAAGACTGACAGAGGCACTTGAAGGGGAGCT 3686  
QY 6259 GCGGCCAGACAGCAGGCCCGCTGAAAGCTTGGCGGGGAGGCGGCCCACTCTCCACACT 6318  
Db 3687 GCGGCCAGACAGCAGGCCCGCTGAAAGCTTGGCGGGGAGGCGGCCCACTCTCCACACT 3746  
QY 6319 GCGGCGGCTGCGAGAGCGGCTCTGTCAGCGCGGCTCTCAAGACCGGCCAGGGGT 6378  
Db 3747 GCGGCGGCTGCGAGAGCGGCTCTGTCAGCGCGGCTCTCAAGACCGGCCAGGGGT 3806  
QY 6379 CAAAGGTACACAGCGGGGTGTCACTCTGGCCAGCAATCATGAGGTATCAACAGGA 6438  
Db 3807 CAAAGGTACACAGCGGGGTGTCACTCTGGCCAGCAATCATGAGGTATCAACAGGA 3866  
QY 6439 CTACACCGGCAACACCAAGCTCAGCGGACCTCTGCGCGCGCGCTCTTACTCTT 6498  
Db 3867 CTACACCGGCAACACCAAGCTCAGCGGACCTCTGCGCGCGCGCTCTTACTCTT 3926  
QY 6499 CCTGGGGGACGCTGCGCGCTCTGAGACCTTCCGCGCGGCAACCAATGACCTTACCTCC 6558  
Db 3927 CCTGGGGGACGCTGCGCGCTCTGAGACCTTCCGCGCGGCAACCAATGACCTTACCTCC 3986  
QY 6559 GCCCGGACCATGTGTGCCCGCGCGCTCTCCCAACAGGAAAGGGGCAAGAGTCT 6618  
Db 3987 GCCCGGACCATGTGTGCCCGCGCGCTCTCCCAACAGGAAAGGGGCAAGAGTCT 4046  
QY 6619 TCCAGAGCCAAACAAAGCTGTGCTTGGGTGTGTGAGAGCGTATTAACCTGTGTC 6678  
Db 4047 TCCAGAGCCAAACAAAGCTGTGCTTGGGTGTGTGAGAGCGTATTAACCTGTGTC 4106  
QY 6679 CCGACCGGAGGGGATGACGAGGACGAGGCACTCCGGAAGTGTGTGATCCCGTGTGTA 6738  
Db 4107 CCGACCGGAGGGGATGACGAGGACGAGGCACTCCGGAAGTGTGTGATCCCGTGTGTA 4166  
QY 6739 CCGGATGAGGAAACAGACGAGCGGACGAGATGGCTCCAACTTCCAGGCAACACAG 6798  
Db 4167 CCGGATGAGGAAACAGACGAGCGGACGAGATGGGCTCCAACTTCCAGGCAACACAG 4226  
QY 6799 CAGCGCGCAAGCTTTTCAAGACTGACCGAGCAACTCCGCAATGCTCAAGTCAA 6858  
Db 4227 CAGCGCGCAAGCTTTTCAAGACTGACCGAGCAACTCCGCAATGCTCAAGTCAA 4286  
QY 6859 GAGGCAAGGATTCACAGAGCTGAAACCCCAACCCGGAATGAGCTGAAATACATAT 6918  
Db 4287 GAGGCAAGGATTCACAGAGCTGAAACCCCAACCCGGAATGAGCTGAAATACATAT 4346  
QY 6919 CAGCAGCTGAGGACGAGATCTTCAATATGCGCGCATCAACCGGAACAGGCTTATGAC 6978  
Db 4347 CAGCAGCTGAGGACGAGATCTTCAATATGCGCGCATCAACCGGAACAGGCTTATGAC 4406  
QY 6979 CTATTAAGGCAAGCGGTGACAGGAACATGCAACCAATGAGGCTGAGGCAATAT 7038  
Db 4407 CTATTAAGGCAAGCGGTGACAGGAACATGCAACCAATGAGGCTGAGGCAATAT 4466  
QY 7039 TAGAAGGCACTCATAGGTAAATATGACAGTGGGAAGTCCCGCGCTGACCGGCA 7098  
Db 4467 TAGAAGGCACTCATAGGTAAATATGACAGTGGGAAGTCCCGCGCTGACCGGCA 4526  
QY 7099 TGCTTTAAACCTCTGATATGCAAGTCCAGGCTGCGCTGTATGCTTAAACCGGTGC 7158  
Db 4527 TGCTTTAAACCTCTGATATGCAAGTCCAGGCTGCGCTGTATGCTTAAACCGGTGC 4586  
QY 7159 TGAACGACGAGAGTACCAACTCACTCTGCGGAGGTGCGGCGGAAAGGCAAGGTCTC 7218  
Db 4587 TGAACGACGAGAGTACCAACTCACTCTGCGGAGGTGCGGCGGAAAGGCAAGGTCTC 4646

QY 7219 TGGCAGACCCAGCAGCCGAAAAAGCAAGTCCCCGGCCCCGGGCTGGCATCTGGGAGCG 7278  
DB 4647 TGGCAGACCCAGCAGCCGAAAAAGCAAGTCCCCGGCCCCGGGCTGGCATCTGGGAGCG 4706  
QY 7279 GCCACCTCTGTCTCTCTAGTGAATCTGGAGGAGAGTGTCAACCGCGGAGCGCTCTAC 7338  
DB 4707 GCCACCTCTGTCTCTCTAGTGAATCTGGAGGAGAGTGTCAACCGCGGAGCGCTCTAC 4766  
QY 7339 CAACCGCGGTGGAGAGCAGGCGCTCGGAGGTTCCAGCCATTCCCTTACAAACC 7398  
DB 4767 CAACCGCGGTGGAGAGCAGGCGCTCGGAGGTTCCAGCCATTCCCTTACAAACC 4826  
QY 7399 CTTGATCATGCGGCTGCAAGCGGAGTGTCAATGCTTCCCAACCCGACCGGCTCCGCG 7458  
DB 4827 CTTGATCATGCGGCTGCAAGCGGAGTGTCAATGCTTCCCAACCCGACCGGCTCCGCG 4886  
QY 7459 GGGCAGCGGCGCTCTGCTGGCCCCCAACCGCTGGAGAGAGAGCCCAAGCTGCT 7518  
DB 4887 GGGCAGCGGCGCTCTGCTGGCCCCCAACCGCTGGAGAGAGAGCCCAAGCTGCT 4946  
QY 7519 CTGCTCGAGTACGAGACCTCTCGCAGAGGAGTGAATCAGAAAGCGGGGGGGGCG 7578  
DB 4947 CTGCTCGAGTACGAGACCTCTCGCAGAGGAGTGAATCAGAAAGCGGGGGGGGCG 5006  
QY 7579 GGGCGGTGERSPERUUGTCAAGTCCAGCGAGCAAGAAAGCGGCTGCAAGAGCGG 7638  
DB 5007 GGGCGG-----TGTCAAGTCCAGCGAGCCAGAGAAAGCGGCTGCAAGAGCGG 5056  
QY 7639 GGGCGGTGCGAATCCCCCAACCAAGAGAGAGCCCTGAGTCCGCTGCTCCATCCA 7698  
DB 5057 GGGCGGTGCGAATCCCCCAACCAAGAGAGAGCCCTGAGTCCGCTGCTCCATCCA 5116  
QY 7699 TCTGTCCGTCCAGAGCGGCGCATCTTGCCTGTAAAGCTTAAGTCAAGTCCCGCCCC 7758  
DB 5117 TCTGTCCGTCCAGAGCGGCGCATCTTGCCTGTAAAGCTTAAGTCAAGTCCCGCCCC 5176  
QY 7759 GGGCGTGGCTGTGAGACCTTACTCAGGGAGTGTTTACTCTGCTGCTGGAGAGGAG 7818  
DB 5177 GGGCGTGGCTGTGAGACCTTACTCAGGGAGTGTTTACTCTGCTGCTGGAGAGGAG 5236  
QY 7819 GAAAGGGCGCGGAGAGGCGACGCGAGGCTGTGGAGCAACACAGGCGCGCACGCG 7878  
DB 5237 GAAAGGGCGCGGAGAGGCGACGCGAGGCTGTGGAGCAACACAGGCGCGCACGCG 5296  
QY 7879 GCCAGAGACCCAAAGAGAGTACCAAGCAAGTCCAGCGCATGCTCCCGGAATGAT 7938  
DB 5297 GCCAGAGACCCAAAGAGAGTACCAAGCAAGTCCAGCGCATGCTCCCGGAATGAT 5356  
QY 7939 TTGGAACCAAGTCTAACTGAGCTGAGCGCCCGCGCGCTCCCTCGGCTCCATCCC 7998  
DB 5357 TTGGAACCAAGTCTAACTGAGCTGAGCGCCCGCGCGCTCCCTCGGCTCCATCCC 5416  
QY 7999 GCTTGGCGCTCTGAGCAGATGAGCGAGGCGCTGTTCAGCGCCAGTGGCTCTCCG 8058  
DB 5417 GCTTGGCGCTCTGAGCAGATGAGCGAGGCGCTGTTCAGCGCCAGTGGCTCTCCG 5476  
QY 8059 TCCCAACAGAGTCCCGCAGCGCAAGGATGTGGAAGAAACAAGTCAAGGCAAGTGGAG 8118  
DB 5477 TCCCAACAGAGTCCCGCAGCGCAAGGATGTGGAAGAAACAAGTCAAGGCAAGTGGAG 5536  
QY 8119 ACAAAGGCGCAGTGGCGCTGGGGGGGAAACGATGCTCCGAGGATGAGTGTTTTTT 8178  
DB 5537 ACAAAGGCGCAGTGGCGCTGGGGGGGAAACGATGCTCCGAGGATGAGTGTTTTTT 5596  
QY 8179 CACACATGTTTCCCGCAGCGGTGGAGAGAAAGGAGATGTAATGATGTTGGTTTAC 8238  
DB 5597 CACACATGTTTCCCGCAGCGGTGGAGAGAAAGGAGATGTAATGATGTTGGTTTAC 5656  
QY 8239 AGGATATTTTTTGAATCCTTCAATGAATTAATTCAGATGTTTACGAAGAGAGACT 8298  
DB 5657 AGGATATTTTTTGAATCCTTCAATGAATTAATTCAGATGTTTACGAAGAGAGACT 5716

QY 8239 ACCGATATCTAGTCTGTGCTTTTGAATCTCTGCTTACCGTTCAAGAGCGGTGTGAG 8358  
DB 5717 ACCGATATCTAGTCTGTGCTTTTGAATCTCTGCTTACCGTTCAAGAGCGGTGTGAG 5776  
QY 8359 GCCGACATCGGTGACCCCATCATCTGAGAGCAAGGGGGGCGGAGCTGCTCGACGC 8418  
DB 5777 GCCGACATCGGTGACCCCATCATCTGAGAGCAAGGGGGGCGGAGCTGCTCGACGC 5836  
QY 8419 CCCGCTGTGCTCTCCCTCCCTCCCTTCCCTTGGGAGAAATGAATGATGATTTCTG 8478  
DB 5837 CCCGCTGTGCTCTCCCTCCCTCCCTTCCCTTGGGAGAAATGAATGATGATTTCTG 5896  
QY 8479 GCCGCAATTTGGCAGGGGTGTGATTTCTGATTTACACAGCTGTTTATTTAAAA 8538  
DB 5897 GCCGCAATTTGGCAGGGGTGTGATTTCTGATTTACACAGCTGTTTATTTAAAA 5956  
QY 8539 GCGAATTTACTCCAAAAAATTTAAAAA 8571  
DB 5957 GCGAATTTACTCCAAAAAATTTAAAAA 5989

## RESULT 6

AF125671

## LOCUS

AF125671 8544 bp mRNA linear ROD 04-APR-1999

## DEFINITION

Mus musculus silencing mediator of retinoic acid and thyroid

## ACCESSION

AF125671

## VERSION

AF125671.1 GI:4559295

## KEYWORDS

Mus musculus (house mouse)

## SOURCE

Mus musculus

## ORGANISM

Mus musculus

## REFERENCE

Chen, J.D.

## AUTHORS

JOURNAL

## JOURNAL

JOURNAL

## FEATURES

source

1..8544

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

1..8544

/gene="Smrte"

160..7548

/gene="Smrte"

/function="transcriptional corepressor"

/note="SMRte"

/codon\_start=1

/product="silencing mediator of retinoic acid and thyroid

hormone receptor extended isoform"

/protein\_id="AAD22972.1"

/db\_xref="GI:4559296"

/translation="MSGSTQPVAVQTRAPRAEPYPHGISYVQIARSHDTVGLSEYOH

HPDYTSLSISGSIIOPORRRPILSERQPSRSOELHLPESRFLPELKGPIBF

TESKRPRLPDLTLPRLPILATQPSGSDLTDRSLAKLPVSPSPRHADLP

ELABRLSKELIQNRDLRVREITWVQIISKLKKQOQLEBAKRPPEKVPSP

PIESHRSLVQIIVDENKKAAEARLLEGAPQVELLYNQPSRQYHENIKINO

MRKLIIFKRRNARKEORPCQYDQLEAMEKVERLENNPRRAKSKVLEY

EKQPEIKQELBQMSRVGSGSAGSAASEHVSSEIIDLSQENLEKQMRQ

LAVIRHVRDQQRKPTINMGDMDPKVKYKIDQVTMMBSQOQDTRKQHPQK

FGLIASFLERKTVAECVLYITLTKENYKSLVRSYRRKSGSQOQOQOQOQOQ

ARSSQEKERKEKEKQDAENEKEKELSKETDTSGBENDEKEAVALSGRK

TANSQGRKGRITRSNANBHEBTAPQOSSLASLMMENSSRWTEBEMETAKGLL  
EHRNWSJALPMVGSKTVSOCKNIFYNYKQULDEILQOHLKMKERKESAPKKKT  
PAAASETAPPAABDEMBASGANSBEELEABASQASGNEVRNENSGARPAVN  
NSDTPSSPSRSEATKQGPCKPTGTEALPAATQPPVPEBEEAALPABSPVPA  
PSPEPSPPAPAPATVDKQBQAPAPAPQTEBDAKQXSBAEITDVQREPEABSE  
PPEVSVDKEETEEBEPEDAKGTEALIVTSEAPLVEBQSAVAATKSSSGATQDS  
DSATCSADBEVEPEGDKRLSPRPSTLTPAGDPRASTSPKDLKQLKORAAI  
PIVTVNHEPREDVTPKPVPPVPPPTQHLQEGDVSOQSGSPGKRSVPVPAEK  
BAEKPAFPAPPTBEGOSYRLSPHAGRLSHSPREVIKTSTRADPLKSTYTPPHPL  
GLHDSAPVLPAPPLISSAGHPVLEBQGLATIQGMSVOLRVHSHBAKMG  
PLTMELPLAVDPKCLSTALAPPEVNASPRASQYPGCRPOLORLTHPRTADVLKGT  
ISRIVGBDSPSRLDRABEDTLPKGNVLYEGKGNVLYSGMSVSCSDKSGSSGP  
PHTABPKRTYDMMGEVGRVTVSASIEGKGNVLYSGMSVSCSDKSGSSGP  
RYVEADQEDYLRBEALIKREGRPPPPPLPTEYTKRPLDPLGLKPHBEPV  
ATYKAGRSIHIEPREELRRTPELPLAPPLKGSITQGPRLKYDSAGSTQKKHDV  
RSIISGRGRFPALHPDLIMADABABRACYESBSLSRSGTSSGAGSITTRGAPVVP  
ELGKQSPUTYEDHGAFTSHLPKGSVATREPTVRLDEGLSLSKASQDRKLSTP  
BIKASPHSTVEHNPPISPYEHILRGVGLVLYRGLPLAVDPSIIRGILPLEAA  
AAVYLRHLAPSTYPLHYRGLYRGLPTALLENQTIINDYITSQMHNHNSAMA  
QADMLRGLSPRESSLALNYAGRGLIDLSQVPHLPVLPPTGPATADILAVLP  
TAPPESSHSSPLSGPPTLAKPTATSSSRERERENDKSLITTTTVEHAP  
WPKPTSSGAGSSRPASHTHQSPISRTQDALQORSVLANTSMKVTVTSVEGT  
PVLNSTSSPVRPAPATPAPATPCPLGVLBSVPLTMEVPLLPKETSIVAPBEXR  
VVGHAFLTKPAPABEPASSPSKSSSRSLAPSSHTALARPXKSLAPHASBPBG  
PVSASDLHREKTOSEKPSIOELERSLGYSAGYSPDVEPISPVSSPLTHDKLS  
KLEBELKHLBEGLHKOQPMKLSABAHLPHLRPLPSOBSPLLOTAGIKGH  
OEVVTLAHLSEVITQDTRHHPQOLSGAPLAPVLPASQVLDLARPDLVLYP  
PHGTPRARSPHSBOGKSPREBKTSVLSSSDALIPVPPREMTBPGHARSAYVLL  
YNDGEGEPRMGSKSPGNTSQPPAPFSQAVOVBHASTVMGLAIIRKALMGYDQWBBP  
LNGOQTEIFNMPLATAGLMTCRQAVOVBHASTVMGLAIIRKALMGYDQWBBP  
LGANAFNPLNASASLPAAMPITADGRDHALTSPGCGKAKVSPRSRRKASVAP  
GLASGRPSPVSVSHSEDPAMNRTPLTNVWEDRPSASTPPVPLMRLQAGVMA  
SPPPPLAAGSGPLAGPHHAMBEPKPLCSQYETLSDBE"

ORIGIN

Query Match 58.8%; Score 5037.2; DB 9; Length 8544;  
Best Local Similarity 78.3%; Pred. No. 0;  
Matches 6770; Conservative 0; Mismatches 1493; Indels 387; Gaps 45;

1 CATGTGGGCTCCACACAGCTTGTGGACAGAGTGAAGGCGCACTGAGCCCGCTAACCC 60  
159 CATGTCAAGATCCACACAGCTGTGGACAGACATGCGGCGCTGAGAGCCCGCTAACCC 218  
61 GCCCCACAGCCTTTCCTACCCAGTGCAGATCGCCGGAAGCAACAGGAGCGTGGGCTGCT 120  
219 ACCCCATGGCATCTCTTACCCGCTGCAATAGCCCGGTCCACACAGGAGCGTGGGCTGCT 278  
121 GGAATACAGACCACTCCCGGACTATGCTCCCACTGTGCGCGGCTCCATCATCCA 180  
279 TAGATACCAACACACCCCGTGACTACACTCAGACTGACCGGTTCCATCATCCA 338  
181 GCCCCACGCGCGAGGCGCTCCCTGCTGTGTGATTCAGACCCGGGAATGAACGTTCCA 240  
339 GCCACAGAGAGAGCGCGCCCTCACTGCTGCAAGTTCAGACCTGGAGTGAAGCGTCTCA 398  
241 GGAAGCTCCACCTGCGCGCGAGATCCCACTCATCTGCGCGAGCTGGAGAGTCAAGAT 300  
399 GGAAGCTCCACCTGCGCGCTGAGTCCCGACAGTTCGCTGAGCTGGAGAGTCAAGAT 458  
301 GGAATTCATTGAAGCAAGCGCCTTCGCTAGAGCTGCTGAGCTGAGCTGCTGAGCC 360  
459 ABAATTCACCGAGAGAGAGCGCCCGCGCTGAGAGCTACACCGGATACCTGTGTGCGCC 518  
361 GTCAACCTCTGCTGCGACAGGCGGACGCTGCGGAGATTGAAGCTTCAACAGAGACCTGAG 420  
519 ATCAACCTCTGCTGCGACAGGCGGACGCTGCGGAGATTGAAGCTTCAACAGAGACCTGAG 578  
421 CCGTGAAGGAGAGCTGGAACCGGCTGTCTCCCGCCAGCCCGCGCACTGAACCTTGAAGCT 480  
579 CTTGGGAGGAGAGCTGGAAGCTGTGTCACTCCAGTCCCGCGCAAGCTGAGCTTGAAGCT 638  
481 GGAAGCTGTGCGCGCAAGCTGTCTCAAGAGAGAGCTGATCCAGAACATTCGAGCCGCT 537

639 AGAGCTGGCGCATCTCGACTGTCCAAAGAGAGCTGATCCAGAACAGATTGAGCCGCT 698  
538 GAGCGAGAGATCAACATGATGTAGAGAGAGATCTCTAAGCTGAAGAAAGAGAGACA 597  
699 GAGCGTGAATACCATGATGAGAGAGAGATCTCCAGCTGAAGAAAGAGAGACA 758  
598 GCTGAGAGAGAGAGCTGCGCAAGCGCGCCGAGCTGGAAGCCCGTGTCAACCGCGCCAT 657  
759 GTTGAAGAGAGAGAGCGCGCAAGCGCGCCGAGCCGGAAGCGGAGCTGTGTCCGACACCAT 818  
658 CGAGTGAAGAGACCGCAGCTGTGTGAGATCATCTTCAAGAGAACCGGAAGAGGCTGA 717  
819 AGATCAAGAGACCGGAGCTGTGTGAGATCATCTTCAAGAGAACCGGAAGAGGCTGA 778  
718 AGCTGACATCGGATCTGGAAGAGCGCTGGGCGCCGAGTGAAGCTGCGCTGTAACAACA 777  
879 AGCGGACACCGGATCTGGAAGAGCGCTGGGCGCCGAGTGAAGCTGCGCTGTAACAACA 938  
778 GCCCTCCAGACCGCGCAGTATCATGAAACATCAAAATTAACAGGCGTCCGAGAA 837  
939 GCGGTGACACACCGCAGTATCATGAAACATCAAAATTAACAGGCGTCCGAGAA 998  
838 GCTAATCTGTACTTCAAGAGAGATACGCTCGGAAACAATGAGACAGAGTTCTG 897  
999 GCTGATCTGTACTTCAAGAGAGATACGCTCGGAAACAATGAGACAGAGTTCTG 1058  
898 CCAAGCGCTTGAACAGCTCATGAGAGCGCTTGAAGAAAGGTGAGAGCGATCGAAGAACAA 957  
1059 CCAAGCGCTTGAACAGCTCATGAGAGCGCTTGAAGAAAGGTGAGAGCGATCGAAGAACAA 1118  
958 CCGCGCGCGCGCGCGCAAGAGAGCAAGGTGCGCGAGTACTAGAAAGCATGTTCCCTGA 1017  
1119 TCCGCAAGAGAGGCGCAAGAGAGCAAGGTGAGAGGATCTAGAGAAACAGTTCCCGGA 1178  
1018 GATCCGCAAGAGAGCGGAGCTGAGAGAGCGCATGACAGACAGGCTGAGCGCAAGCGCGAG 1077  
1179 GATCCGCAAGAGAGCGGAGCTGAGAGAGCGCATGACAGACAGGCTGAGCGCGAG 1238  
1078 TGGGCTGTCAATGTGGCGCGCGCGAGAGACAGAGGTGTCAGAGATCATGATGATGAGCC 1137  
1239 TGGGCTGTCAATGTGGCGCGCGCGAGAGAGCGCATGAGAGGTTTCTGAGATCATGATGATG 1298  
1138 CTCAAGAGAGAGAGCTGAGAGAGAGAGCGCAGCTGAGCGCTGATCCGCGCATGCT 1197  
1299 GTTGAAGAGAGAGAGCTGAGAGAGAGAGCGCAGCTGAGCGCTGATCC--GCCATGTT 1356  
1198 GTACAGAGCTGACACAGAGCGCATCAAGTTTCAATCAATGAACGCGCTTATGCGCAACC 1257  
1357 GTACAGAGCTGACACAGAGCGCATCAAGTTTCAATCAATGAACGCGCTTATGCGCAACC 1415  
1258 CATGAAGGTGTCAAGAGAGAGAGCGCGCATGAACTATGTGAGGTGAGAGAGAGAGAGC 1317  
1416 CATGAAGGTGTCAAGAGAGAGAGCGCGCATGAACTATGTGAGGTGAGAGAGAGAGAGC 1475  
1318 CTTCCGAGAGAGAGTCAATGACAGATCCAAAGAACTTGTGCTGATGCAATCTCTGGA 1377  
1476 CTTCCGAGAGAGTCAATGACAGATCCAAAGAACTTGTGCTGATGCAATCTCTGGA 1535  
1378 GAGAGAGAGAGTGTGAGTGTGCTCTTATTAATACTGACTTGAAGAAATGAAGATTA 1437  
1536 GAGAGAGAGAGTGTGAGTGTGCTCTTATTAATACTGACTTGAAGAAATGAAGATTA 1595  
1438 TAAAGAGCTGTGAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1497  
1596 CAAAGAGCTGTGAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1643  
1498 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1557  
1644 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1703  
1558 GAAAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1617  
1704 GAAAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1763

OY	1618	GAACGACAAGGAAGA	CTTCTTCAAGAGAGA	CAGACGAC	CTTCAAGGAGAGACA	CAACGA	1677
Db	1764	GAACGAGAAGGAAGA	ACTCAGCAAGAGAGA	CAGACGAC	CTTCTGCGAGAGACA	CAACGA	1823
OY	1678	CGAGAAGAGGCTGTGG	CTCCCAAGGCGGAAA	CTGCGCAACAGCGAGGA	AGACGCA	1737	
Db	1824	TGAGAAAGAGGCGTGG	CTCCAAAGGCGGAAA	CTGCGCAACAGCGAGGA	AGACGCGCGCA	1883	
OY	1738	AGGCGCATCA	CCCCGTATATGAGG	CCAAACGAGGAGGCGCATCA	CCCCCA	1797	
Db	1884	AGGCGGTATCA	CGCGCTCATATGCCAA	GAGGGCCAA	CCATATAGAGAGACGCCA	CCCCCA	1943
OY	1798	GCAGAGCCCGAG	CTGGCTCCCATGAG	CTGATATGAGATTCTG	CTGCGAACAAGAAAGA	1857	
Db	1944	GCAAGATT	CAGAGGCTGGCTTCAT	TGAGATGAACAAGAGTT	CTCGCTGACATGAGAAAGA	2003	
OY	1858	AATGGAAACAG	CGAAGAAAGGCTCTG	TGAAACAGGCGGCAAC	TGCTGGGCATTCGCGC	1917	
Db	2004	GATGGAGACAG	AAAGAAAGGCTTCT	TGAAACATGAGAGAA	CTGCTATAGCCATTGCGC	2063	
OY	1918	GATGTGGGCT	CCAAGCTGTGTG	CAGTGTAAAGACTTCT	CAACTCAAGAAAG	1977	
Db	2064	CATGTGGGCT	CCAAGACGTGTCC	CAATGTAAAGACTTCT	CAACTCAAGAAAG	2123	
OY	1978	GCAGAACTCGAT	GATCTTGTGACAG	ACCACTTGAAGATGAGAA	AGAGAGAAAGC	2037	
Db	2124	GCAGAACTCGA	CGAAATCTTACAG	ACCACTAAAGATGAGAA	AGAGAGAAAGC	2183	
OY	2038	GCAGAGAAAGAA	GAAGACGCGCGCGCGC	CAAGCGAGAGAGGCTG	CAATCCGCGCGT	2097	
Db	2184	TCGAGAGAAAGAA	GAAGACCCCAAGCTG	CGCGAGCGAGAGACAGCTT	CCCACTTCG	2243	
OY	2098	GCTGAGGATGAG	AGATGAGGCGT	CGGCGTGA	CGGAAATGAGAGAGATG	2157	
Db	2244	CGTGAAGACGA	AGATGGAAGCAT	CAGCGCAAGTCC	CAATGAGAAAGCTG	2303	
OY	2158	GAGGCTGAAC	CTTACATGCTCT	CTGGGAATGAGGTGCC	CAAG--GGGAAATG	2214	
Db	2304	GAGGCAAAAC	CTCACAGGCTCT	CTGGAAATGAGGTTCC	CAGAGTTGGAGAGTGA	2363	
OY	2215	CCGAGCACTGT	CAACAAGCTCAG	ACCGAGAGCAATCC	CTCTCTCACTAGAGC	2274	
Db	2364	CCGAGCTGT	CAACAACAGTCT	GATATCTGAGAGTCC	ATCCCGCTTCAAGAC	2423	
OY	2275	CGCCAAAGACA	CAAGGCAGAATGG	CGCCCAAGCCCC	CACACCTTGGCGCG	2334	
Db	2424	CACGAAGAC	AC-----TGGGCT	TAACCCCACTG	AGACTGAAAGATTGCGCGTGC	2474	
OY	2335	AACCCCAAGCC	CAACCCACCA	CCAGAGACATCC	CGGCGCCCCATTGAGCCCA	2394	
Db	2475	CACCCAGCC	ACTGTTCTCTCT	CAAGAAACCGG	CAAGCCCCCTTGTGAGCCCC	2534	
OY	2395	GAGCTCTGA	AGCACCAGGACCC	CTACAGCCCA	CCCCCATCGCCTTGCAC	2454	
Db	2535	AGTCCCTAT	TGCCA---GTGG	CCACATCCCA	AGGCTTCCCATAC	2591	
OY	2455	TCCTCTGTGT	CCCCAAGAGAGAG	AGAGAGAGAC	CCGACAGCGCCCCCA	2514	
Db	2592	CCCGGCT	CTGTGGACAAG	AGATGAACAAG	CCCCGCTGCTCA	2651	
OY	2515	GGAGGGGAG	AGAGAGAGCCCC	CGCGCGCTGAG-----	GAGCTGGC	2556	
Db	2652	AGATGCCA	AGAGAGAGAGTCT	GAGAGCCGAGAGAT	TCATGTGGAAACCA	2711	
OY	2557	AGTGAACA	CAAGGAAAGCC	CGAGAGCCGTCA	AGACGATGCA	2616	
Db	2712	CGAGGCT	CTTGAAGAGCC	CCGAGAGATGTAA	AGATACCA	2771	
OY	2617	GGGCGCGC	---CAAGG	CAAGACCGGAGGCT	CGCTGACGCGCCAGGCGGCGCT	2673	
Db	2772	AGAGCT	GAAGCAAAAGCCAA	GGGCAAGAGGCT	CAATGAACTGTGTG	2831	

OY	2674	CAAGCAGAGAAAGAGAGGCGCGAGCGGCGAGGCGCACCACTGCGCAAGAGCTCGGCGCG	2733
Db	2832	TAAG---GTGAGAGAGGCTGTGACGAAGCGAGCTGTGACCAAGGTTTCCAGCTCAGGTGC	2888
OY	2734	CCCCCAGAGCAGCGACTCCAGTGCCTACCTGACAGTGCAGACGAGGTGATGAGCCGAGGG	2793
Db	2889	CACCCAGAGACGTGACCTCCAGTGCACCTGCAGTGCAGATGCCGATGAGGTGAGCCAACTCGAAGG	2948
OY	2794	CGCGCAGCAAGAACCGGCTGCTTGTCCCAAGGCCCAAGCTTCTTCACTCCGACTGGCAGACC	2853
Db	2949	AGGTGACAAAGGAGGAGGTGCTGTGTCACCAAAGGCCAGCTCTTCACTCCGCTGGAGATTC	3008
OY	2854	CCGGGCGCAATGCTTCACTCCCAAGAACCACTGACACTGAAACAGCTGAAAGCAGGAGCGGC	2913
Db	3009	CCGGGCGCAGTACTCTCGCCCCAGAGCGCTGAGACTGTAAACAGCTGAAAGCAGGAGAGC	3068
OY	2914	TGCATCTCCCCCTCCAGGTACCAAAAGTCCATGTAGCCCCCGCGGAGAGCAGCTCC	2973
Db	3069	CGCATCTCCCCCTTATC---GTACCAAAAGTCCATGTAGCCCCCGCGGAGAGCAGTACC	3125
OY	2974	CACCAAGCCAGCTTCCCCAGCCCCCAGCCGACCGCAAAACCTGCAAGCGGAGAGCGACGC	3033
Db	3126	CCCAAAACAGTTCCTCCCTGTGCTTCCACCGCAGACACTTACAGCCAAAGGATGACGT	3185
OY	3034	CCCTCAGACAGCTGTGACAGACCCCCCGGGGCAAGAGAGAGAGCCCGGACACCCCGCGCA	3093
Db	3186	GTCTCAGACAGTCCGGAGAAAGTCCACGTGGCAAGTCCCGAGCTCCAGTGTCTTCTTGCGA	3245
OY	3094	CAAGAGAGGC-----CTTGCACGCCAGGCGCCAGAACTGCT	3130
Db	3246	GAAAGAGGACAGAGAAACCGCATTTCTTCCGGCTTTTCCCACTGAGGGCCAAAGCTACCG	3305
OY	3131	GGGAGCCCCCTTGTGTAATTCGAGCTGACCTTCCCGGTGCCCTTCCCGTGCCTCCGCTGAGGTATC	3190
Db	3306	ACTGAGCCCCCAGCTGTGATC-----CGGCTGCTTCCATCTTCCACGGAGAGTATC	3360
OY	3191	AAGGCTTCCCCCGCATGCGCCCCGAGACCTCTAGACTTCTTCTTCAAGCTCCACCTGGTCAACCA	3250
Db	3361	AAGCTTTCACAGCGCG-----TGACCTCTCTTCTTCTTACACACCTCCCGGTACACCG	3414
OY	3251	CTGCCCCCTGGAGCCTCCATGACACTGACCAGGCGCGTCTGTGCGGCGCCACCACTATCTC	3310
Db	3415	CTGCTCTTGGAGCTTCCACGATAGTGGCCCCGCGCTGTGCAGS--TCCCCCATCTCT	3471
OY	3371	GCCATCTTCCCAAGGAATGTGGTTCAGCTCAGCTCCGTACTCAGAGCATTCGCAAGGCC	3430
Db	3532	GCCATCTTCCCAAGGGAATGATCAGTTCAGCTTCTGTGTCTTCACTCAGAGCATTCGCAAG---	3588
OY	3431	CCGGTGGGCGCTGTGACATAGGGGCTGGCGCTGGCCATGGAACCCAAAAAGCTGGACCC	3490
Db	3589	CCCATGGGCGCTTCCACCATGAGAGCTGCGCCCTTGGCGTGGACCTTAAAGAGCTG-----	3642
OY	3491	TTTCAGCGAGTGAAGCAGAGCAGTGTCCCAACGGGGCCAGGCTGGGCCACCGAGAGGC	3550
Db	3643	-----	3642
OY	3551	CTGGGGGTGCCACAGCCCAAGAGAGCGTCCGTGTGAGAGGAGCAGTCTTGGGTCACTT	3610
Db	3643	-----GGGACAGCACT--GGCTTCGCC	3662
OY	3611	CCGGGCGGAAGCATACCAAAAGGATTTCCAGACACAGGGGTGCCCTGGACAGGCGCATC	3670
Db	3663	ACCAAGTAAAGCATCACAAAGGGCTTCCAGTACCCGGGCTGGACAGCG-----CCCC	3715
OY	3671	ACATACCGCGGCTTCATCACCAACGACCGCAGCTGACGTCCTGTACAAAGGCAACCATC	3730
Db	3716	AGCTACAGAGGCTTATCACCCAC--GGAAGCGCAGAACGTCCTTACAAAGGGATCATTC	3774
OY	3731	ACCAGATCATCGGCGAGACAGCCCGAGTTCCTTGGACCGCGGCGGAGAGACAGCTTG	3790



```

Db      3775 AGCAGATCGTCGTGAGGACAGCCCAAGTCGCTTGAACGGGACAGAGAGACACCTTG 3834
Qy      3791 CCCAAGGCGCACGCTCATCTACGAAGGCAAGAGGCGCCACGCTTTGCTCTTATGAGGGTGC 3850
Db      3835 CCCAAGGCGCATCTCATCTATGAGGGCAAGAGGCGCCACGCTCTTCTCTTGAAGGTGAT 3894
Qy      3851 ATGTCTGTGACCGCAGTGTCTCAAGAGGACGGGCAAGAGCAGCTCAGGACCCCCCATGAG 3910
Db      3895 ATGTCTGTGTCAAGTGTCTTAAAGAGAGATGAAAGAGACGCTCGGGCCACCCCATGAG 3954
Qy      3911 ACGGCGCGCCCAAGCGCACCTTATGACATGATGAGGGCGCGTGGGCAAGGCCATCTGC 3970
Db      3955 ACTGCGCGCCCTAAAGCAGCCTATGACATGATGAGGGCCGTGTGAGGCAAGCTGTACC 4014
Qy      3971 TCGAGCAGATCGAAGTCTCATAGGGCCGTGCCATCTCCGCGGAGGCAACAGCCCCAC 4030
Db      4015 TCGCCAGCATTAAGGAGCTCATAGGGCCGCGCATCTTGAGCAGC-----ACAGCCCC 4068
Qy      4031 CACCTCAAGAGGACGACACATCCGCGGGTCCATCAACAAGGAGATCCCTCGGTCTAC 4090
Db      4069 CACCTCAAGAGAGCATCATCTCGAGGCTCATACGAAAGCATCCGAGGTCTTAT 4128
Qy      4091 GTGAGAGCACAAGAGGACTACTGCGTCGGAGGCGCAAGCTCTTAAAGCGGAGGGCAGC 4150
Db      4129 GTGAGGCGCAGAGGACTTACGGCGGGAGGCCAAGCTTTGAAGGAGAAAGGAGCA 4188
Qy      4151 CCTCCGCGCCCAACCGCCCTCAAGGAGCTGACGGAAGCTCTTAAAGCGGAGGGCAGC 4204
Db      4189 CCACCTCCCCCACAACCACTCTGGAGCTGACGTGAAGCTTCAAGGCCCGCCCTTGAGC 4248
Qy      4205 ----CTGAGCGCCCTGAAGCTGAAGCGGGCCATGAGGGGCTGGTGGCACAAGTGAAGAG 4261
Db      4249 CCTCTGGGTCTCTTGAAGCTGAAGCGGACTCAAGAGGTTGGTAGCAATTGAAGAGAG 4308
Qy      4262 GCGGCGCGCTCCATCCATGAGATCCGCGGAGAGCTGCGGCAACGCGCGAGCTTGCC 4321
Db      4309 GCGGCGCGCTTATCATATGAGATCCGAGAGAGAGCTGCGGCAACCTGAGCTTACC 4368
Qy      4322 CTGGCCCCCGCGCGCTCAAGAGAGGCTCATACGAGGGGACCCCGCTCAAGTACGAC 4381
Db      4369 CTGGCAACAAGGCTCTGAAGAGGGTTCATCACCAAGGGCAACCCCACTCAAGTACGAC 4428
Qy      4382 ACCGCGGCGTCAACACTGGCTCAAAAAGCAGAGGTACGCTCCCTCAATCGGCAACCCC 4441
Db      4429 TCTGGGAGCACTTCACTGAGCAACAAAGAACAGAGTGCGCTCAATCACTGGCAAGCCC 4488
Qy      4442 GGCAGGACGTTCCACCCGTGACCCGCTGATGTATGAGCGACCGCGGACACTTGAA 4501
Db      4489 GGCAGGCTTTTCCCTGCGCTGACCCGCTGACATATGCTGACCCCGGGACACTGGAG 4548
Qy      4502 CGTGCCTGCTACAGAGAGAGCTGAAGAGCGGCGCAAGGACCGGCAAGCACTCGGAGGC 4561
Db      4549 CGTGCCTGCTATGAAGAGTGAAGAGCGGCTCAAGGAGCAAGCAGTGGTGAAGGGGCG 4608
Qy      4562 TCCATTGGCGGCGGGCGCCCGGTCAATTGGCCGTAAGCTGGTGAAGCGCGGCAAGACCCC 4621
Db      4609 TCCATACACGTGGGGCTCAAGTGTGTGCTGAATGGGCAAGGCAAGCAAGCCCA 4668
Qy      4622 CTGACCTTATGAGAGCACAAGGGGCAACCTTTGCGGCACTCCCAAGAGTTGCGCCGTG 4681
Db      4669 CTGACTTAAAGAGACACAAGGGGCAACCTTCAACAAGTCACTGCAAGTGGCTCCCTGTG 4728
Qy      4682 ACCATGCGGAGCGCACGCGCGCGCTGACAGAGGGGAGGCTTTGCTCAGCAAGGCACTCC 4741
Db      4729 ACCACAGAGGAGCGCACGCGCCTTCAAGAGGAGGAGCTCTCTATCAAGCAAGGGGTCC 4788
Qy      4742 CAGGACGAAAGCTGACGTGACGCGCTGAGATGCGCAAGTCCCGGCAAGCAAGCAAGCTG 4801
Db      4789 CAGGACGGAAGCTGACATCAACCCCGGAGATGCGCAAGTCCCAACAGCAAGCTGTG 4848
Qy      4802 CCCGAGCACCAACCCCACTCTGCGCTATGAGCACTGCTTGCGGGCGTGAAGTGGC 4861

```

---

```

Db      4849 CCCGAGCACCAACCTTCACCCCATCTCCCTATGAGCACTTGTCCGGGGCGTGAAGTGT 4908
Qy      4862 GTGACCTGTATGCGAGGCCACATCCCGCTGGCCTTGAACCCCACTTCATACCCCGGCG 4921
Db      4909 GTGACCTGTATCGGTGTACATCCATGAGCCTTGAACCCCACTTCATACCCCGAGGG 4968
Qy      4922 ATCCCTGTGGA---CGAGCGGTGCTTACTACCTGCCCCGACAGCACTGGCCCCAACCCC 4978
Db      4969 ATCCCTGTGGAACACAGCCGAGGCTTACTACTGCCCCGGGACATTTGGCCCCAGCCCC 5028
Qy      4979 ACCTACCCGCACTGTATCCCAACCTTACTCATCGCGGCTACCCCAACGCGCGGCTG 5038
Db      5029 ACCTACCCCAACCTGTATCCCACTTACTCATCGCGGCTACCCCAACGCGCGGCTG 5088
Qy      5039 GAGAACCGGAGCACTATCAATGACTATCATCACTTCGAGAGATGACCAACAACAG 5098
Db      5089 GAGAACCGGAGCACTATCAATGACTATCATCACTTCGAGAGATGACCAACAACGCT 5148
Qy      5099 GCCACCGCAATGGCCCAAGGAGCTGATATGCTGAGGGGCTCTCGCCCGGAGTCTCG 5158
Db      5149 GCTCCGCGCATGGCCCAAGGCTGTACATGTGAGGGGTGTGTCAACGAGAGATCTCG 5208
Qy      5159 CTGGCACTAATTAATTCGCTGGGCTCCCGAGGATATGAGCTGTGTCCAAAGTCCAC 5218
Db      5209 CTGGCCCTCAATTAATTCGCTGGGCTCCAGAGGATATGAGCTGTGTCCAAAGTCCAC 5268
Qy      5219 CTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5278
Db      5269 CTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5328
Qy      5279 TACCTCCCAACCGCGCCGAGGCTTCAAGAGCGGCAAGAGAGTCTCCCACTCTGCCA 5338
Db      5329 TACCTCCCACTGTGCGGCCCAACCTTCAAGAGGCGCAAGTATGATCAACGCTGTGCCA 5388
Qy      5339 GAGAGTCCAAACAATTGAACAAAACCAACCAACGATCTCTGTCCAGCGGAGCGAGAC 5398
Db      5389 GAGAGCGCCCACTCACTAAGTAAACCACTGACCACTTTCATCGAGGGGAGACG---- 5444
Qy      5399 CGGAGTCAAGAGCGGGACCGGGATCGGGAGCGGGAAAGTCTCTTCACTGTCCACACAG 5458
Db      5445 -----GGAAGTGAAGCGGGAACAGAAAGTCAATCTCACTGTACCACT 5490
Qy      5459 ACGGTGAAGCAGCAGCACTTGAAGAGCTGTGAAGAGCAGAGCAGGAGCAGCAGCGC 5518
Db      5491 ACAGTGAAGATGACCACTTGAAGAGCTGTGAAGAGCAGAGCAGGAGCGGAGCTGGGGC 5550
Qy      5519 AGCAGCGCGGGGGGTGGGGGACAGCAAGCCGCGCTTCCACTCCATGCCCCAAG 5578
Db      5551 AGCAGCGG-----CCCGGCTCCCAACCCACAG 5580
Qy      5579 CACTGCGCCATCTCCCTGAGACCCAGATGCCCTTCAAGCAAGACCCAGTGTCTTAC 5638
Db      5581 CACTGCGCCATCTCCCGCGAGACCGCTTGAAGAGGCGCTTGAAGAGGCGCAGTGTGAC 5640
Qy      5639 AACAGAGCATGAAGGATATCATCAACGCTGTGAAGCCAGCAAGGCCACGCTTGAAG 5698
Db      5641 AACAGAGCATGAAGGAGCTGTCACTTCTGTGAACCCGGCAACGCTCTGTGAAG 5700
Qy      5699 TCACTTCAACTCTTCACTCACTGTGCGCAGCTGCAATTCCTCACTTCCACCTCACTG 5758
Db      5701 TCACTTCACTCTTCTGCTGTGCGCCAGCTGCAATTCCTCACTTCCACCTCACTG 5760
Qy      5759 CCACTGAGGCGCACTCTGATGGGGTGTACCTTACCTTCAATGAAGCCGCTTTGCTGCC 5818
Db      5761 CCACTTGTGGGCACTTGAAGGGGTGTACCTTACCTTATGAAGCGCTCTGTATACC 5820
Qy      5819 AAGAGGCGCCCGGGTGCAGGCGGAGCGGAGCGGAGCAGAGCAAGCCGCAATGCTTC 5878
Db      5821 AAGAGAGCTCTCGGGTGCAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCTTC 5880
Qy      5879 CTGCGCAAGCGCCCAAGCGGCTCGGGCTGAGAGCGGCTCTCTCCCAAGAGGCTCG 5938
Db      5881 CTACACAAACCCCGGCGG-----GAGAGCGGCTCTCTCAACCCAGCAAGGCTCC 5931

```





Db 8066 TCCAGTCTCCAAA---GGCCACCCAGCCCAAGAGCTGGAGCCCATGACAGCAGGTGG 8122  
Qy 8115 GCGGACAAAAGGCGCAGAGTGGCGGCTGGGGGGAACGAGATGCTCCAGACCTGAGCTGTT 8174  
Db 8123 GTGACACAAAGGGGCTGG-----CTGGGGACCGAGTCTTGACGAGACTGACCGTT 8174  
Qy 8175 TTTTCAACATCGTTGCGCGGAGCGGTGGGAAGAAAGCAGATGTAAATGATGTTGGT 8234  
Db 8175 TCCGGGCTGTGTCTGTGGCAAGG---GAGGGAAGGACGTTAAATG--GTGTTGGC 8228  
Qy 8235 TTTACAGGTATATTTTGTATCTTCAATGATTAATTCAGATGTTTACGCAAGAGG 8294  
Db 8229 TTACAGGGATATTTTGTATCTTCAATGATTAATTCAGAGCTTCACACAGAGAGG 8288  
Qy 8295 ACTTACCCAGTATTTACGTGTGTGCTTTGTATCTTGTACCTTACGCTTAAAGAGCGTG 8354  
Db 8289 ACTGCGCC--GTGTTTCTCCGCTGTGC--TTTGGTCTTACTTACTGTTTCAAGGCAAGT 8346  
Qy 8355 GCAAGCCGACAGTGGTGAACCCCATCACTGCAAGACCAAGGGGCGGGAGCTGCTGTC 8414  
Db 8347 GC---CAGCAAGGCGGTGGCCACATACGAGGACTTGGGGGTCAAGGGGCTCCGGCAC 8403  
Qy 8415 AGCCCGCGTGTCTCTCCCTCCCTCCCTTCTTGGGAGATGAATTGATGCTGATTC 8474  
Db 8404 AGCGCACTGTG---CCCTTCCCAACCTTCACTTCAAGCAAAAGACTTGATCGTATTC 8460  
Qy 8475 TGTGGCGCATTTTGGGAGGGTGTGTATCTGTCACTTACACAGCTGCTTAATTA 8534  
Db 8461 TGTGGCGCTCTGTGCA---CGGCGGCACTTGTGCTTACATGTTGTTCAATTA 8517  
Qy 8535 AAAAGCGAAT 8544  
Db 8518 AAAAGCAAT 8527

RESULT 7 8388 bp mRNA linear ROD 20-MAR-1999  
LOCUS AF113001  
DEFINITION Mus musculus silencing mediator of retinoic acid and thyroid  
hormone receptor alpha mRNA, complete cds.  
ACCESSION AF113001  
VERSION AF113001.1 GI:4454547  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
REFERENCE 1. Mus musculus  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Muridae; Murinae; Mus.  
1 (bases 1 to 8388)  
Ordentlich,P., Downes,M., Xie,W., Genin,A., Spinner,N.B. and  
Evans,R.M.  
TITLE Unique forms of human and mouse nuclear receptor corepressor SMRT  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (6), 2639-2644 (1999)  
PUBMED 10077563  
2 (bases 1 to 8388)  
Downes M.R., Ordentlich,P. and Evans,R.M.  
REFERENCE Direct Submission  
TITLE Submitted (11-DEC-1998) Gene Expression Laboratory, The Salk  
JOURNAL Institute for Biological Studies, 10010 North Torrey Pines Road, La  
Jolla, CA 92037, USA  
FEATURES  
source location/Qualifiers  
1..8388  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/tissue\_type="spleen; brain"  
626..8047  
/function="transcriptional co-repressor"  
/note="SMART alpha"  
/codon\_start=1  
/product="silencing mediator of retinoic acid and thyroid  
hormone receptor alpha"

CDS  
1..8388  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/tissue\_type="spleen; brain"  
626..8047  
/function="transcriptional co-repressor"  
/note="SMART alpha"  
/codon\_start=1  
/product="silencing mediator of retinoic acid and thyroid  
hormone receptor alpha"

## misc\_feature

## ORIGIN

Query Match 58.2%; Score 4981.6; DB 9; Length 8388;  
Best Local Similarity 79.9%; Pred. No. 0;  
Matches 6351; Conservative 0; Mismatches 1265; Indels 331; Gaps 28;  
Qy 1 CATGTGGGCTCCACAGAGCTTGGACAGAGCTGAGAGGCGCACTGAGCCCGCTAACCC 60  
Db 625 CATGTGAGATCAACAGAGCTTGGACAGAGCTGAGAGGCGCTGCTGAGCCCGCTAACCC 684  
Qy 61 GCCCAGAGCCCTTCTTACCCAGTCCAGATGCGCCGGAACGACACAGAGCTGCGGCTCCT 120  
Db 685 ACCCCATGGCATCTCTACCCCGGTGCAATAGCCGGTCCCAAGAGAGTGGGCTGCT 744  
Qy 121 GAGATACAGACCACTCCCGGACTAGCGCTCCCACTGTGCGCGGCTCCATCATCA 180  
Db 745 TGAGTACCAACACACACCCCGGACTACCACTGATCAACCGGTTCATCATCA 804  
Qy 181 GCCCAGGCGGAGAGGCGCTCCCTGCTGTGAGTTCCAGCCCGGGAATGAACGGTCCA 240  
Db 805 GCCACAGAGAGAGGCGGCGCTCTGCTGAGAGTTCCAGAGCTGGAAGTGAACGGTCTCA 864  
Qy 241 GAGAGCTCACTGCGGCGGAGAGTCCCACTATCTGCGCGAGAGCTGGGGAATCAAGAT 300  
Db 865 GAGAGCTCACTGCGGCGGAGAGTCCCACTATCTGCGCGAGAGCTGGGGAATCAAGAT 924

QY 301 GGAGTTCAATTGAAAGCAAGCGCTTCGAGTAGCTGTGCTCGTGAACCCCTGCTGCGAC 360  
DB AGAATTTCACCGAGCAAGCGCCCGCTGAGGCTACTACCGGATACCTGTCTGCGGCC 984  
QY 361 GTCAACCCCTGTGCGGCAAGCGGCAAGCTGTGCGGAACTTGAAAGCTTCAACAGAACCTTGA 420  
DB ATCAACCCCTGTGCGGCAAGCTGTGCGGCAAGCTGTGAAAGCTTGAACCTTGAACAGAACCTTGA 1044  
QY 421 CCTGACGGGCAAGCTGTGAACCGGATGTCTCCCGCAAGCCCGGCAACATGAACCTGAGCT 480  
DB 1045 CTTGCAAGGCAAGCTGTGAAGCTGTGTCTCACTTCCAGTCCCGGCAAGCTGTGAACCTTGAAGCT 1104  
QY 481 GGAAGTGTGCTGCGCAACCGCTGTCTCAAGAGGAGCTGATCCAGAACATGGAACCGGTGGA 540  
DB 1105 AGAGCTGGGGCCATCTCGACTGTCCAGAGGAGGAGCTGATCCAGAACATGGAACCGGTGGA 1164  
QY 541 CCGAAGATCACATGATGTAAGACAGACATCTTCTAAGCTGAAAGAAAGAACAGCAACGCT 600  
DB 1165 CCGTGAAGATCACATGATGTAAGACAGACATCTTCTAAGCTGAAAGAAAGAACAGCAACGCT 1224  
QY 601 GGAGAGGAGGCTGTGCAAGCGCCCGAGGCTGAGAAAGCCGATGACCGCGCCCATGCA 660  
DB 1225 GGAGAGGAGGCTGTGCAAGCGCCCGAGGCTGTGATGCTGTGCTGCAACCATGGA 1284  
QY 661 GTGGAAGCACCGCAGCCTGTGTGCAATCTTACGAGAAACCGGAAGAGGCTGAAAGC 720  
DB 1285 ATCAAGACCGAAGCTGTGTCAAGATCATCTACGATGAGAACCGGAAGAAAGCGGAAGC 1344  
QY 721 TGCAATCGGATTTGGAAGGCTGTGCGGCGCCCAAGTGTGAGCTGTGCTGTACACACGCG 780  
DB 1345 CGCAACCGGATCTTGAAGGCTGTGCGGCGCCCAAGTGTGAGCTGTGCTGTACACACGCG 1404  
QY 781 CTCCGACACCGGCGAGTATCATGAGAACATCAAAATTAACAGGCGATGCGGAAGAGCT 840  
DB 1405 GTCTTACACCGCGAGTACATGAAACATCAAAATTAACAGGCGATGCGGAAGAGCT 1464  
QY 841 AATCTTGTACTTCAAGAGAGGAAATCAAGCTGTGAGAAACATGAGAAACAGAAATTCTGCA 900  
DB 1465 GATCTTGTACTTCAAGAGAGGAAACAGCGCGCAAGCAAGTGGAAACAGCGCTTCTGCA 1524  
QY 901 GCGCTATGACCGAGCTCATGAGAGGCTTGGAAAAAAGTGTGAGCGGATCGAAAAACACCC 960  
DB 1525 GCGCTATGACCGAGCTCATGAGAGGCTGTGAGAAAGGATGAGCCCATGAGAAACATCC 1584  
QY 961 GCGCGCGCGGCGCAAGAGAGGAGTGTGCGAGATCTTACGAAAAAGCATTCCTGAGAT 1020  
DB 1585 GCGAAGAGGCGCAAGAGAGGAGTGTGCGAGATCTTACGAAAAAGCATTCCTGAGAT 1644  
QY 1021 CCGCAAGCAGCGAGCTGTGCAAGAGCGCATGCAAGCAGGCTGTGAGCGGCGGAGTGTG 1080  
DB 1645 CCGCAAGCAGCGGAGCTGTGCAAGAGCGCATGCAAGCAGGCTGTGAGCGGCGGAGTGTG 1704  
QY 1081 GCTGTTCATGTGTGCGCGCGCGCAAGAGACAGAGTGTCAAGATCATGTGAGCTGTG 1140  
DB 1705 GCTGTTCATGTGTGCGCGCGCGCAAGAGAGGTTTCTGAGATCATGTGAGCTGTG 1764  
QY 1141 AGAGCAGAGGAACCTGAGGAACAGATGTGCGCAGCTGTGCGGATCCCGCCCATGCTGA 1200  
DB 1765 TGAGCAGAGGAACCTGAGGAACAGATGTGCGCAGCTGTGCGGATCCCGCCCATGCTGA 1824  
QY 1201 CGACGCTGACCGAGCGCATCAAGTTCTCAACATGAAACCGGCTTATGTGCGACCCCAT 1260  
DB 1825 CGACGCTGACCGAGAGATCAAGTTCTCAACATGAAATGGAATCATGATGAGACCCCAT 1884  
QY 1261 GAAAGTGTACAAAGACCGCAGGCTCATGAAACATGTGAGTGTGACGAGAGAGAAAGACTT 1320  
DB 1885 GAAAGTGTACAAAGACCGTCAAGTTTACAAACATGTGAGTGTGACGAGAGAGAAAGCACTT 1944  
QY 1321 CCGGGAAGATTCAAGCAGCATCCCAAGAACTTGTGCGGATCGATCATCTCTGGAAG 1380  
DB 1945 CCGTGAAGATTATGACAGCACCTTGAAGACTTGTGCGGATTCCTCTGGAAG 2004

QY 1361 GAAGCAGTGTGCTGAGTGTGCTCTCTTATTAATTAATTAATTAATTAATTAATTAATTA 1440  
DB 2005 AAAGCGGTGCTGAGTGTGTCTCTTATTAATTAATTAATTAATTAATTAATTAATTA 2064  
QY 1441 GAGCTGTGAGACGGAAGCTATCGGCGCGGGAAGAGGCGGAGGAGCAACAGCAGCA 1500  
DB 2065 GAGCTGTGAGGCGGAAGCTATCGGCGCGGGAAGAGGCGGAGGAGCAACAGCAGCA 2124  
QY 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
DB 2125 ACAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2172  
QY 1561 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
DB 2173 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2232  
QY 1621 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
DB 2233 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2292  
QY 1681 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
DB 2293 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2352  
QY 1741 CCGCATCACCGGCTCATGTGCTTATGAGGCGCAACAGAGAGAGAGAGAGAGAGAGAG 1800  
DB 2353 CCGTATCACCGGCTCATGTGCGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2412  
QY 1801 GAGCGCGAGGCTGTGCTCATGTGAGTGTGATGAGATCTGTGCTGTGAGAGAGAG 1860  
DB 2413 AAGTTCAAGAGCTGTGCTCATGTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 2472  
QY 1861 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920  
DB 2473 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2532  
QY 1921 GGTGGGCTTCAAGAGCTGTGTGTGCAAGTGTGAAGACTTCTTCACTTCACTTCAAGAGAGCA 1980  
DB 2533 GGTGGGCTTCAAGAGCTGTGTGTGCAAGTGTGAAGACTTCTTCACTTCACTTCAAGAGAGCA 2592  
QY 1981 GAACTCGATGAGATCTTGTGAGCAGCAGCAAGCTGAGATGAGAGAGAGAGAGAGAGAG 2040  
DB 2593 GAACTCGATGAGATCTTGTGAGCAGCAGCAAGCTGAGATGAGAGAGAGAGAGAGAGAG 2652  
QY 2041 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
DB 2653 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2712  
QY 2101 GAGAGATGAGAGATGAGAGGCTGTGAGGCTGTGAGGCGGAATGAGAGAGAGATGTGAGAG 2160  
DB 2713 TGAGAGAGAGAGATGAGAGGATGAGGCGCAAGTGTGCAATGAGAGAGAGAGAGAGAG 2772  
QY 2161 GGTGTGAAGCTTCAATGTGCTGTGAGAGATGAGTGTGCGCAAGAGAGAGAGAGAGAGAG 2217  
DB 2773 GGTGTGAAGCTTCAATGTGCTGTGAGAGATGAGTGTGCGCAAGAGAGAGAGAGAGAGAG 2832  
QY 2218 AGCAGCTGTCAACAAACAGCTCAGACACGAGAGAGATCCCTCTCTCTCACTGTAGAGCGCG 2277  
DB 2833 AGCAGCTGTCAACAAACAGCTCAGATCTGATATCTGAGAGTGTCCCATCTCCGCTTCAAGAGCAT 2892  
QY 2278 CAAGAGCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2337  
DB 2893 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2943  
QY 2338 CCGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2397  
DB 2944 CCGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3003  
QY 2398 CTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2457  
DB 3004 CCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3059  
QY 2458 TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2517

Db 3060 GGGTACTGTGAGCAAGGATGAAACAAGAGCCCGGCTGTCTCAAGTCCCAACAAGAGA 3119  
QY 2518 GGGGGAGAGCAGAGCCCGCGGCTGAG-----GAGCTGGCAG 2558  
Db 3120 TGCAGAGAGCAAGAGTCTGAGGCCAGAGGATCGATGTGGGAAAAACCAAGAGAGCCG 3179  
QY 2559 TGGACACAGAGGAAAGCCGAGAGGCCGCTCAAGAGCAGATGCAAGAGAAAGCCAGAGAG 2618  
Db 3180 AGGCCCTGTAGAGAGCCCGAGAGAGTGTAAAGATGACCAACAAGAGAGAGACCGAGAG 3239  
QY 2619 GGGCGGC---CAAGGCAAGAGACGCGAGAGCCGCTGAGGCAAGCGCGAGAGAGGAGCTCA 2675  
Db 3240 AGCTGAAAGCAAAAGCCAAAGGGGACAGAGGCAATTGAAACTGTGTGAGGACCACTTA 3299  
QY 2676 AGGCAGAGAAAGAGAGGCGGAGGCGGAGGGCCACCACTGCGCAAGAGTCCGGGCGCC 2735  
Db 3300 AG---GTGAGGAGGCTGTGAGCAAGGCAAGCTGTGACCAAGGCTTCCAGCTCAGTGTCCA 3356  
QY 2736 CCCAGGACAGCGACTCCAGTGTCTACCTGAGTGCAGACGAGTGGATGAGGCGGAGGAG 2795  
Db 3357 CCCAGGACAGTGAATTCATGATGCACTTGAGTGCAGATGAGTGAACGAACCGAAGAG 3416  
QY 2796 GCGACAAAGAACCGGCTGTGTCTCCCAAGGCGCAGCTCTCAACCCGACTGGCAAGCCG 2855  
Db 3417 GTGACAAAGGGCAGGCTGTGTGTCAAAAGGCCAGGCTCTCAACCCGGCTGGAGATCCCG 3476  
QY 2856 GGGCCCAATGCTTCACCCCAAGAGCCACTGGAAGCTGAAAGAGCTGAAAGCAGGCGCTG 2915  
Db 3477 GGGGCCAGTACCTCGCCCAAGAGCCGCTGGAAGCTGAAAGAGCTGAAAGCAGCGAGCG 3536  
QY 2916 CCATCCCCCCCCATCCAGGTCAACCAAGTCCATGAGGCCCCCGGGAGAGACGAGCTCCCA 2975  
Db 3537 CCATCCCCCTTATCCAGGTCAACCAAGTCCATGAGGCCCCCGGGAGAGACAGTACCCC 3596  
QY 2976 CCAAGCAGACTCCCCCAGGCCCAACGCGCAACCGCAAAACCTGACCGGAGAGAGCGCC 3035  
Db 3597 CAAAGCAGATTCCTCCCTGTGCTCCACCAAGAGAGACCTTACAGCAGAGGAGTGAAGTGA 3656  
QY 3036 CTCAGCAGCTTGGCAGAGAGCCCGGGAGGCAAGAGAGAGCCGCGACACCCCGCGGACA 3095  
Db 3657 CTCAGCAGTCTGGAGAGAGTCCACGTGGCAAGTCCGCGAGCCAGTGTGCTCTCGCGAGA 3716  
QY 3096 AGGA-----GGCTTGGAGCCGAGGCGCCAGAGAGCTGCTG 3131  
Db 3717 AAGAGGAGAGAAACCGCAATCTTTCGCGCTTCCCACTGAGGAGCCCAAGCTACCA 3776  
QY 3132 GGGACCCCCCTTGTGAGACTTCCGAGCTGCGCTTCCCGTACCCCGGTGAGGTATCA 3191  
Db 3777 CTGAGCCCCCAAGTGTGATGAGGAGCTGCTTCCCTTCCCTCAACGAGAGGTATCA 3836  
QY 3192 AGGCTCTCCGAGATGCGGAGACCCCTCAGCCTTCTTACGCTCCACCTGTGACCCAC 3251  
Db 3837 AAGCTTCCCCACACGCGCGCTGACCCCTGTGCTTCTTACACACCCCGGATCACCGC 3896  
QY 3252 TGGCTTGGGCTCTCATGACATGCCCCGCGCTGTGCGCGGCGCACCACTCATTTCCA 3311  
Db 3897 TGCCTTGGGCTCTCAACGATGAGCCGCGCGCTGTGCAAG---TCCCCCATCTCTCA 3953  
QY 3312 ACCCGCTCCCCCTCATCTCTGTGCAAGACACCGAGCTCTGAGAGGCAAAATAGTG 3371  
Db 3954 ACCCCCCACCCCTCATCTCTCTGTGCAAGATCCCGCGTACTTGAAGAGGAGCTGGGTG 4013  
QY 3372 CCATCTCC---CAAGAAATGTGAGTCAAGCTCCAGCTCCGTACTCAGAGATGCCAAG 3428  
Db 4014 CCATCTCCCAAGAGGGAGTGTCAAGTTCAGCTTGTGTGCTCACTCAGAGATGCCAAG 4073  
QY 3429 CCGCGGTGGGCTCTGTACACATGGGGCTGCGCTGCGCAATGACCCCAAAAAGTGGAC 3488  
Db 4074 CCCCCATGGGCTCTCAACATGAGGAGCTGCCCTTTCGCTGAGCCTTAAGAAAGCTG--- 4129  
QY 3489 CCTTACGCGAGATGAAAGAGAGAGAGCTGTCCACGAGGAGCAGGCTGGGACACCGAGA 3548

Db 4130 ----- 4129  
QY 3549 GCTTGGGGGTGCCACAGCCAGAGAGCGCTCGGTGTGAGAGGGAAGCTGTGGAGCTAG 3608  
Db 4130 -----GGAACAGCACTGGGCTCG 4148  
QY 3609 TTCCGGGCGGAAGCATCACCAAGGATTTCCAGCAACGCGGTGCGCTTGGACAGCGCA 3668  
Db 4149 CACAGTGGAGATCACCAAGGGGCTCCAGATACCGGG-----CTGACAGACGGCC 4202  
QY 3669 TCACATACCGCGCTTCATCAACCAAGCAAGCTCAGCTGACGTCTGTATCAAGGGACCA 3728  
Db 4203 CCAGCTTACAGAGGCTTATCACCAAGCAAGCGCCCAAGCGTCTTCAAGAGGTATCA 4262  
QY 3729 TCACAGAGATCATTCGGGAGAGCAGCCGAGTGTGAGTGTGAGACGCGCGCGGAGAGAGCC 3788  
Db 4263 TCAGCAGATCGTGGTGAAGACAGCCAAAGTGCCTTACCGGGGACAGAGAGACACC 4322  
QY 3789 TGCACAAAGGACATGCTATCTACGAAGCAAGAGGACAGCTTGTCTTATGAGGAGT 3848  
Db 4323 TGCACAAAGGACATGCTATCTATGAGGACAAAGAGCCAGCTCTATCTTATGAGAGT 4382  
QY 3849 GCATGTCTGTACCCAGTCTCCAGAGAGACGCGCAAGAGAGCTACGAGACCCCGCATG 3908  
Db 4383 GTATGTCCGTGTACAGTGTCTTAAAGAGATGAAAGAGAGCTCGGCGCCACCCCATG 4442  
QY 3909 AAGCGGCGCCCGCAAGCGCACTTATGACATGAGAGGCGCGGTGGCAGAGCCATCT 3968  
Db 4443 AAGCTGCGCCCTTAAACACACTTATGACATGAGAGGCGCTGTAAAGAGAGTGTCA 4502  
QY 3969 CCTCAGCCAGATCGAAGGTCTCATGTGGCGGTGSCATTCGCGCGAGAGCAACAGCCCC 4028  
Db 4503 CTTAGCAGCATGATGAGGATCTATGAGGCGCGGCACTCTCTAGAGCA-----CAGCC 4556  
QY 4029 ACCACTTAAAGAGCAGCAACATCCGCGGTTCATCAACAAGGATCCTCGGTCT 4088  
Db 4557 CCACCTCAAGAGAGCAGCATCATCCAGGCTCCATCAACGAGGATCCCGAGGTCT 4616  
QY 4089 AGGTGAAGCACAAGAGAGCTACGTGCTGGGAGGCGCAAGCTCTTAAAGCGGAGGCA 4148  
Db 4617 ATGTGAAGCGAGAGAGCTACTTACGCGGAGGCGCAAGCTCTTGAAGCGAAGGGA 4676  
QY 4149 CGCTCCGCGCCCAAGGCTTACAGGAGCTGACCGAGGCTTACAGAGCGCAGAGCC---- 4204  
Db 4677 CACACACACCCCAACACCACTCCGGAACCTGACGTGAGACTTACAGAGCCCGGCTCG 4736  
QY 4205 -----CTGGGCCCCCTGAAGCTGAAGCCGCGCCATGAGGCGCTGTGGCAAGTAAAG 4259  
Db 4737 ACCCTTGGGTCCCTGAAGCTGAAGCGCATCAAGAGGTGTGTAGCACTGTAAAG 4796  
QY 4260 AAGCGGCGGCTTCATCATGAGATCCGCGAGAGAGCTGCGGACACGCGGAGCTGC 4319  
Db 4797 AGGCGGCGGCTTATCATGAGATCCGAGAGAGAGCTGCGGCGCACACTGAGCTAC 4856  
QY 4320 CCTGGCGCGGCGGCGCTCAAGAGAGGCTCATATCAAGAGGAGACCCGCTCAATAG 4379  
Db 4857 CCTGGACACAGGCTCTGAAGAGAGGTTCAATCAACAGGAGACCCCACTCAATAG 4916  
QY 4380 ACACCGGCGGTTCACACTGTGCTCAAAAAGCAAGACGTAGCTCTCATCGAGCC 4439  
Db 4917 ACTCTGGGCAACCTTCATGAGCAACAAAGAAACAGACGTGGGCTCATCATGAGAGCC 4976  
QY 4440 CCGGCGGAGGTTCCACCCGTGCAACCGGTGAGTGTATGCGGACGCGCGGCACTGG 4499  
Db 4977 CCGGCGGCTTTCCTGTGCGACCCGCTGACATAATGTGCTGACGCGGCGCACTGG 5036  
QY 4500 AAGGTGCTGTACAGAGAGAGCTGAAGAGCGGCAAGGGAACCGGCAAGCTCGGGG 4559  
Db 5037 AGGTGCTGTATGAAGAGTGTGAAGAGCGGTCAGAGACCAAGCATGTGTGAGAGG 5096  
QY 4560 GCTTCATTGCGGCGGCGCGCTGATTTGTGCTTGAAGTGTGATAGCGCGAGAGCC 4619  
Db 5097 GCTTCATCAACAGTGGGCTTCAAGTGTGTGCTTGAAGTGTGGCAAGGACCGCAAGGCC 5156

QY 4620 CCTGACCTATGAGGACCAAGGGGACCCCTTTGCGGCCACTCTCCACAGAGGTGCGCG 4679  
DB 5157 CACTACCTTAAAGAACACACGAGGACCTTTCAACAGTCACTGCACAGTGGCTCCCG 5216  
QY 4680 TGAACATGCGGAGCGCCACGCGCGCTGACAGAGGGACGCTTTCTGTCAGCAAGGCAT 4739  
DB 5217 TGAACACAGAGGAGCGCCACGCGCTTCAAGAAAGGCGCTTCTATTCACAAAGGCGT 5276  
QY 4740 CCACAGACCGAAAGGAGCTGACGCGCTGTAAGATCGCCAAAGTCCCGCACAGACCG 4799  
DB 5277 CCAAGACCGAGAGCTGACATTTACACCCGGAGATGCGCAAGTCCCAACAGCATG 5336  
QY 4800 TGCAGAGACCAACCAACCCCATCTGACCTATGAGACCTGTCTTGGGGCTGAGTG 4859  
DB 5337 TGCCGAGAGACCTCTACCCCATCTCCCTATGAGCATTTGCTCGGGGCGTGACTG 5396  
QY 4860 GCGTGAACCTGTATGCGACGCCAATCCCTGAGCTTGCACCCCATCTCATACCCCGG 4919  
DB 5397 GTGTGAACCTGTACCGTGGTCACATCCCATTTGACCCCACTTCATACCCCGAG 5456  
QY 4920 GATCCCTCTGGA---CGAAGCGCTGCGCTTACTGCGCCGACACCTGGGCCCAAC 4976  
DB 5457 GATCCCTCTGAGAGACGCGACGCTTACTGCTGCGCCGACCTTGGGCCCGAGC 5516  
QY 4977 CCACCTACCCGACCTGTACCAACCTTACTCATCGCGCTAACCCGACAGCGGCGC 5036  
DB 5517 CCACCTACCCACACCTGTACCAACCTTACTCATCGCGGCTTACCTGACAGGGGCGCC 5576  
QY 5037 TGAAGAACCGGACAGCATCATCATATGACTACATCACTTCGACAGATGACACACAA 5096  
DB 5577 TGAAGAACCGGACAGCATCATCATATGACTACATCACTTCGACAGATGACACCAAG 5636  
QY 5097 CGGCGACCGGACATGCGCGACGAGTATGCTGAGGGGCGCTCGCGCCCGAGTCT 5156  
DB 5637 CTGCTCTCGACATGCGCGACGCGTGTGACATGCTGAGGGGCTGTGACCGCAGAGTCT 5696  
QY 5157 CGCTGGACATCAACTACGCTGCGGGTCCCGAGGCGATCATGACCTGTCCAAAGTGCAAC 5216  
DB 5697 CGCTGGCGCTTAATTAATGCGCTGCGCCCAAGAGCATTAATGACCTGTCCAAAGTGCAAC 5756  
QY 5217 ACCTGCGTGTGCTGTGCGCCGACACAGGACACCCGACCAACCGCATGAGACCGCTTG 5276  
DB 5757 ACCTGCGCGTGTGCTGTGCGCCGACCAAGCGACCGCTGCAACCGCATGAGACCGCTTG 5816  
QY 5277 CCTACTCCCGACCGGCGCCGACCGCTTGAAGACCGGACAGAGTCCCACTCTCC 5336  
DB 5817 CCTACTCTCCCGACGCGCGCCGACCGCTTGAAGACCGGACAGAGTCAACGCTGTCC 5876  
QY 5337 CAGAGGTCCAAACACTTGAACAAACCAACCAACGCTCTGTCGAGCGGAGCGAG 5396  
DB 5877 CAGAGGCGCCCACTACCTTAAGTAAACAACTGCGACATCTTCAATGAGGCGG----- 5929  
QY 5397 ACCGGATCGAGAGCGGAGACCGGATCGGAGCGGAGAAAGTCCATCTCACGTCCACA 5456  
DB 5930 -----GAACGGAAACGTAGCGGGAAGAGAACAAAGTCCATCTCACTGTAACA 5978  
QY 5457 CGACGCTGAGACACGACCTTGTGAGACCTGTGTAAGAGACAGAGCGGACAGAGCG 5516  
DB 5979 CTACAGTGAAGATGACCATCTGTGAGACCTGTGTAAGAGACAGAGCGGAGCTGGG 6038  
QY 5517 GAGAGAGCGGCGGGGGTGGGGGAGAGAGCGCGCGCGCTCCCATCTCCCATGCGCAC 5576  
DB 6039 GAGAGCGCGGCG-----CGGCTTCCACACCCAC 6068  
QY 5577 AGCACTCGCCCATCTCTCCCTCGGACCCAGAGATGCCCTTCAGACAGAGACCCAGTGTCTTC 5636  
DB 6069 AGCACTCGCCCATCTCTCCCGGAGCCCAAGAGCGCTTGAAGAGAGCGCCAGTGTCTGC 6128  
QY 5637 ACACACAGAGCATGAAAGGTATCATACCGCTGTGAGCGCCAGCAAGCCACGCTCTG- 5695  
DB 6129 ACACACAGAGCATGAAAGGTGTGATCACTCGTGAACCCGCGCACGCGCACGCTCTGA 6188

QY 5696 -----AGGTCACTCTCAACCTCTCAACCCGTTGCGCCAGCTGCGACATTTCCACTG 5747  
DB 6189 GATGGGCCAGGTCACCTTCCACTCTTCTGCGCTGTCCCGCCAGCTGCGACATTTCCACTG 6248  
QY 5748 CCACCACTGGCCACTGGGGCGGACCCCTGATAGGGGTCTACCTTACCTCATAGAGCCG 5807  
DB 6249 CCACCACTGGCCACTTGTGTGACCTTTGAAGGGGTCTACCTTACCTCATAGAGCCG 6308  
QY 5808 TCTTGTGCGCCAGAGAGGCGCGCGGGGTGCGCGCGGACAGAGCGCGCGGACAGCACCG 5867  
DB 6309 TCTGTTAACCAAGAGACCTGTGGGGTGTGCGCGCGGAGCGGCGCGGGGTGAGCGCTG 6368  
QY 5868 GCATGCTCTCTCGCCACAGCGCCCGCTCGGGCTGAGAGCGCGCTCTCTCCCA 5927  
DB 6369 GCATGCTCTTCTTAACAAACCCCGGCGG-----GAGCGCGCGCTCTTCAACCA 6419  
QY 5928 GCAAGGGCTCGAGAGCGCGCGCGCTTATGCGCTGTCTGAGGACGCGCACATCGCGC 5987  
DB 6420 GCAAGAGCTCGAGAGCGCGGATCTTACGACCGCCGAGCTCAAGCGCACAGCATGTGCC 6479  
QY 5988 GCACCCCTGCAAGAACCTGCACTTCAACGCGCAGCGCGGACCGCGCGCACCTG 6047  
DB 6480 GCACCCCGCAAGAACCTTGAACCCCAACATGCGAGTCCGAGCCGCGCGG---CCCA 6536  
QY 6048 CCTGGCTCTGAGACCGGACCGGAAAGACTCAAGGTAAACCTTTTCATCCAGAAC 6107  
DB 6537 CCTGGCTCTGAGATGCAACGAGAAAGACTCAAGGTAAACCTTTTCATCCAGAAAT 6596  
QY 6108 TGAACCTCGTCTGTGGTTAACAC---GGCAGAGCTTACAGCGCGGAGGGGTGAGC 6164  
DB 6597 TGAACCTCGTCTGTGGTTAACAGAGTGAAGCTGAGCTTACAGCGCGGAGGGGTGAGC 6656  
QY 6165 CCGTGGCTGTGAGCTCACCAAGTGTGACCCAGCAAGAGGGGTCTCCCAAGCATG 6224  
DB 6657 CCATAGCGCGGTGAGCTTCCCGAGCTGACCCAGACAGAGAGGGGTCTTCCAAACCTGCG 6716  
QY 6225 AAGAGCTGACAGAGGCACTGTGAGGGGAGCTGCGCGCCAGACGCGCGCGCTG 6284  
DB 6717 AAGAGCTGAGAGAGGCACTTGTGAAGGGAGCTGCGGACACAGAGCGCGCGCCATGA 6776  
QY 6285 AGCTTGGGGGAGGCGCGCCACCTTCCGACCTGCGCGCGCTGTGAGAGCGAGCCT 6344  
DB 6777 AGCTGAGGCGGAGGCTGCCATCTTCCACATGTGCGGCACTGCGCGAGCGCGCT 6836  
QY 6345 CGTTCAGCGCTGTCCAGACCGCGCGGAGGTCAAGGTCAACGAGGGGTGTCAACC 6404  
DB 6837 CATTCAGCGCTCTTCCAGAGTCCCGAGGATCAAGAGGTCAACGAGGGGTGTCAACC 6896  
QY 6405 TGGCCAGACATCATGAGGTGATCACAAGAGCTTACACCCGAGCAACCCACAGAC 6464  
DB 6897 TGGCTACAGACATCAGCGAGGTGATTAAGCAGAGCTTACACCGGCGACACCGGAGCAGC 6956  
QY 6465 TGAAGCACCTTGGCGCGCGCGCTTACTCTTCCCTGGGGGCGAGCTGCGCGCTG 6524  
DB 6957 TCAAGTGGCGCGCTTCCGCGCTTCTTACTCTTTCGCGAGCGAGCTGCTGTCTG 7016  
QY 6525 ACTTCGCGCGCGCACCATGAGCTTACTCTTCCGCGCGCGGACCATGATGCGCGCGC 6584  
DB 7017 ATCTTGGCGCGCGCACCATGAGCTTACTCTTCCACCGCGCGGACCATGAGCGCGCGC 7076  
QY 6585 GTGCTCCCGCCACAGCAAGGAGGCAAGGTCTTCAAGAGCGCAACAGAGTGTCT 6644  
DB 7077 GGGGATCCCGCCACAGTGAAGGGGCAAAAGGTCCCGAGAACCCAGCAAAACATCGGTCC 7136  
QY 6645 TGGGTGTGTGAGAGCGGTATTGAACCTGTGTCCCAACCGGAGGCGATGACGAGCGAG 6704  
DB 7137 TGGGACAGAGCGAGGATGCACTTGAAGCTGTGTCCCAACAGAGGCGATGATGAGCGAG 7196  
QY 6705 GGCACCTCCCGAGTCTGTGTACCGGCTGTGTAACGGGATGGGGAAAGAGAGCGCA 6764  
DB 7197 GACATGCTGTGAGCACTGTGTACCACTGTGTATCGAGACGGGAAACGGGCGAGCC-- 7254  
QY 6765 GCAGGATGGGCTCAAGTCTCCAGGCAACACAGCGAGCGCGCAAGCTTCTTCAAGCAAGC 6824



[illegible]



HTIATARPANLAPHASPPRAPATASDLHNEKQSRFSTQELHLSLGSYHSGAGY  
SPOVDEPISPSVSLTHDQSGKRPYLEEKLHSEBETODYDNRHPOOLSAAPLAPH  
RLPDESOPSSLOTLQAPKIGKORVLAQHSITVTDYDNRHPOOLSAAPLAPH  
SFPASCPVADLRPEBDLPLPDGTGTARSGPHBEGSGSRPESKTSVLSESDIA  
EVPSPGMEPEGARSTAVLYLRGBOEGPMLGSGSGNTSOPFTSLTSSSADA  
MYSEKQIKNTKHNHNRPNYJIOGPETFNMDATGALMTCSOAYOEHAESTN  
MGELAITRKLMGTIDOWEPPPLGANAFLPLNASLPLAAMPITTDJGSDIALTS  
PGGGAKKAVSGPSSRAKSPAGLAGSGRPVSSVSHSGGONNRPLTNRWEDR  
SSASSTFPYNNPLIMRLQAGVMSPPPGPLAAGSGPLAGPHAMWDEBKPLT  
"ISDSE"

11

Query Match	51.1%	Score 4378.2;	DB 9;	Length 7465;
Best Local Similarity	79.2%;	Pred. No. 0;		
Matches 5700;	Conservative	0;	Mismatches 1164;	Indels 334;
			Gaps	29;

QY	750	CCGAGGTGGACCTGCGCGCTGTACCAACACGAGCCCTCCGACACCCGGGAGTATCATGAGAAC	809
Db	454	CCCGGTCCCAACGCGCTTGTACCAACGAGCCGTCTGCACACGCGATGACATGAAACA	513
QY	810	TCAAAATTAACACGAGCGCATGCGGAAAGAGCTAATCTTGTACTTCAAGAGGAATCACG	869
Db	514	TCAAAATTAACAGAGCGCATGCGGAAAGAGCTGATCTTGTATCTTTAAGGGAGAACACG	573
QY	870	CTCGGAAACAATGGAAGCAAGTTCTGCGACCGCTATGACACAGTCAATGAAGCCTTGG	929
Db	574	CGCCGACGCAATGGGAAACAGCGCTTCTGCACAGCGTATGACACAGTCAATGGAAGCGTGG	633
QY	930	AAAAAAGGTGAGCGCATCGAAAAACAACCCGCGCGCGGCGCAAGAGAGCAAGGTGC	989
Db	634	AGAAAGAGGTGAGCGCATATAGAAACAATCCGCGAAGGAGGGCCAGAGAGCAAGGTGA	693
QY	990	GCGAGTACTACGAAAAGCATTTCCCTGAGATCCGCAAGCGCGAGCTGCAGAGCGCA	1049
Db	694	GGGAGTACTAGAAACAGTTCCCGGAGATCCGCAACACCGGGAGCTGCAGAGCGCA	753
QY	1050	TGCAGAGCAGGSGTGGGCGAGCGGGGCAATGGGSGCTGTCAATGCGGCGCGCGACGAGC	1109
Db	754	TGCAGAGCAGGSGTGGGCCAGCGGTGCAATGGGCTTCCATGTGCGCTGCCGAGTGAAC	813
QY	1110	ACGAGGTGTCAAGATCATTCGATGGCTCTGAGAGCAGAGAACCTTGAGAGACAGATGC	1169
Db	814	ATGAGGTTCTTGAGATCATTCATGGCTTGTCTGAGCAGAGAACCTTGAGAGACAGATGC	873
QY	1170	GCCAGCTTGCGCGTGTATCCCGCCCATGCTGTACAGACGCTGACACAGCGCATCAAGTTCA	1229
Db	874	GCCAGCTTGCGCGTGTATCCCGCCCATGTTGTACAGACGCGACAGCAGAGAGATCAAGTTCA	933
QY	1230	TCAACATGAAACGGGCTTATGGCCGACCCCAATGAAGGTGTAACAAGAACGGCAGGTCATGA	1289
Db	934	TCAACATGAAATGAGACTCATGTGATGACCCCAATGMAAGTCTTACAAGAACCTGTACGTTACA	993
QY	1290	ACATGTGAGTGAAGCAGGAGAAAGAGACCTTCCGGAGAAATTCAATGACATCCCAAGA	1349
Db	994	ACATGTGAGGAGACAGGAGAGAGAGACCTTCCGTAAGAAATTTATGACAGACCTTAAGA	1053
QY	1350	ACTTTGGCTGTATGCGATCATTTCCGTGAGAGAGAAAGACATGTGCTGAGTGCCTCTATT	1409
Db	1054	ACTTTGGCTGTATGCTCTCATTTCTGTGAGAGAAAGACGCTGCTGAGTGTCTCTATT	1113
QY	1410	ACTACCTGACTAAGAGAGATGAGAACTATAAGAGCTTGATGAGACGAGCTATCGAGCC	1469
Db	1114	ACTACCTGACCAAGAGAAATGAAAATTACAAAGAGCTTGATGAGCGAGCTATCGAGCC	1173
QY	1470	GCGGCAAGGCCAGACGACCAACAACGACGACGACGACGACGACGACGACGACGACG	1529
Db	1174	GTCGCAAGAGC-----CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	1221
QY	1530	AGCCCATGCCCCGAGCAGCCAGGAGGAGAAATGATCAAGAGAGAAAGAAAGAGCGG	1589
Db	1222	AGCAGATGCGACGAGCAGCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGCGG	1281
QY	1590	AGAAAGAGAGAGAAAGCCGAGAGGTGAGAAACAACAAGAAAGCCTTCTCAAGAGAGA	1649

Db	1282	ACAAAGAGAGAGAGAGAGAGATTCGGAGAAACGAGAAAGAGAGCTGAGGCTCCAAAGGCGCGCA	1341
QY	1650	CAGACGACACCTCAGGGGAGAGCAACGACGAAAGAGAGCTGTGGCTCCAAAGGCGCGCA	1709
Db	1342	CAGACGACACCTTCTGGCGAGGAGCAACGATGAGAAAGAGCGCTGTGGCTCCAAAGGCGCGCA	1401
QY	1710	AAACTGCCCAACAGCCAGGAGAAAGCGAAAGGCGCATACCCCGCTCAATGGCTTAATGAG	1769
Db	1402	AAACTGGCCAAACAGCCAAAGCGCGCAAGGCGCATACCGCTGTCAATGGCCCAACAGG	1461
QY	1770	CCAACAGGAGAGAGGCGCATCACCCCCAGCGAGACGCGCGAGCTGGCTCCATGAGAGCTGA	1829
Db	1462	CCAAACCATGAGGAGACAGCCACCCCAAGCGAAATTCAGAGCTGGCTTCCATGAGAGATGA	1521
QY	1830	ATGAGAGTCTTCGCTGAGCAAGAAAGAAATGGAACAGCCCAAGAAAAGTTCCTTGAGAC	1889
Db	1522	ACGAGAGTTCCTCGCTGAGCTGAGGAAAGATGAGGACAGCCAAAGAAAAGGCTCTTGAGAC	1581
QY	1890	ACGCGCCGCACTGGTCCGCGCATCGCCCGGATGGTGGGCTCCAGACTGTGTCCGACGTGA	1949
Db	1582	ATGGAGAGAACTGGTTCAGCCATTCGCCGCAATGGTGGGCTCCAAACCGGTGTCCACGTGA	1641
QY	1950	AGAACTTCATCTTCAACTCAAGAAAGAGGCGAGAACTCGATGAACTCTTGAGAGAGACA	2009
Db	1642	AGAACTTCTACTTCAACTCAAGAAAGAGGCGAACTTGAGAGAAATCTTTCAGACAGACA	1701
QY	2010	AGCTGAAGATGAGAAAGAGAGAAACGCGCGGAGAAAGAAAGAAAAGCGCCGCGCGCG	2069
Db	1702	AGCTAAAGATGAGAAAGAGAGAAACGCTCGGAGAGAAAGAAAGAAAGAACCCCAAGCTCGG	1761
QY	2070	CCAGCGAGAGAGCTGCATTCGCCCGCCGTGTGAGAGATGAGAGATGAGAGCGTTCGGGCG	2129
Db	1762	CGACCGAGAGACAGCGCTTCCCACTTCGCCGTGAGGCGAAGAGATGAAAGCACTCAAGCG	1821
QY	2130	TGACCGGAAATGAGAGAGAGATGGTGGAGAGGCTGAAGCCCTTACATGCGCTGTGGAGATG	2189
Db	1822	CAATGTCGCAATGAGAGAAAGAGCTGCGGAGAGAGGAGAAAGCTTCAGAGGCTCTTGGAGATG	1881
QY	2190	AGGTGCCGACA---GGGAAATGCAGTGGGCGCAGCACTGTCAACAAACAGCTCAGACACG	2246
Db	1882	AGGTTCGCAAGATTTGGGAGATGCAGTGGCCAGCTGTGTCAAAACAGCTCTGATACTG	1941
QY	2247	AGAGCATCCCTCTCTCAACTGAGGCGCGCAAGACACAGGCGAGATGGGCTCCACG	2306
Db	1942	AGAGTGTCCCATCCCGCGTTCAAGAAACCAAGAAAGACAC-----TGGGCTTAAC	1992
QY	2307	CCCCAGCACCTTGGGGGCGGACGGGCGACCCCCAGGCCCCACCCCAACCCCAAGAGGA	2366
Db	1993	CCACTGGCACTGAAGCAATGCGCGCTGCGCAACCCAGCCACCTGTTCTCTTCAAGAAAC	2052
QY	2367	CATCCCGGGCCCCCATTTGAGGCCCAACCCCGGCTTTGAAGCCACCGAGCCCTTACGCCCC	2426
Db	2053	CGGAGATGAGCCCCCTGTGAGCCCTTCCCAATGCCCTTGAATGCCAGTG-----CCCAACATCC	2108
QY	2427	CACCAAGCACCCCAATGCGCTTGTGCACCTCTCTGTGTGTCCTCCAGAGAGAGAGAGG	2486
Db	2109	CCAGAGGCTTTCATCACTGTCGCGCACCCCGGCTTACTGTGACAAAGATGAACAAGAG	2168
QY	2487	AGGAGACCGCAGACGCGCCCCCAATGAGAGAGGGGAGAGACAGAAACCCCCCGGCGCTG	2546
Db	2169	CCCCGGCTGTCTCCAGCTCCCAAGCAAGAGATGCAAGGAGCAAAATCTGAGGCCAGAG	2228
QY	2547	AG-----GAGCTGGCAGTGAACAAGGAAAGGCGAGAGACCGCTC	2587
Db	2229	AGATTCATGTGGGAAAAACAGAGAGAGCCGAGGCTTCTGAGAGGCCCCGAGAAAGTGA	2288
QY	2588	AAGAGCGAGTGCACGAGAGAGCCGAGAGAGGCGC---GGCCAAGGCGAAAGACGCGAG	2644
Db	2289	AAGAGTGACCAACAAAGAGAGAGACCGGAGAGAGCGCTGAAAGCAAAAGCCAGAGGCAAGAG	2348
QY	2645	GCGCCTGAGGCGACGCGCGAGGGGCGCTCAAGCAAGAGAAAGAGAGCGCGGAGCGGCG	2704

Db 2349 GCCATTGAACCTGTGTCTGAGGACCACTTAAG---GTGAGAAAGGCTGTGTGACAAAGCA 2405  
Qy 2705 AGGAGCAACCACTGCGAAGAGCTCGGCGCGCCCCAGAGACAGAGTCCAGTGTACTCTGC 2764  
Db 2406 GCTGTGACCAAGGGTTCAGCTCTCAGTGTCCACCCAGAGACAGTGTACTCAGTGTCCACTGC 2465  
Qy 2765 AGTGCAGACGAGGTGTAGAGCGGAGGCGGCGACAAAGAACCGGTGTGTCTCCCAAG 2824  
Db 2466 AGTCCCATGTAGGTGTGACGAACCCGAAAGAGGTGACAAAGGCGAGCTGTCTGTCAAG 2525  
Qy 2825 CCCAGCCCTCTCAACCCGACCTGTGAGACCCCGGCGCAATGTCTACTACCCCAAGACTG 2884  
Db 2526 CCCAGCCCTCTCAACCCGCTGTGAGATCTCCGGGCAATCTCTGCCCCAAGACCGCTG 2585  
Qy 2885 GACCTGAAGCAGTGTGAAGCAGCGGCTGCACTCCCGCCATCCAGGTCAACAAAGTC 2944  
Db 2586 GACCTGAAGCAGTGTGAAGCAGCGGCTGCACTCCCGCCATCTCCCGCCCAAGACCGCTG 2642  
Qy 2945 CATTAGGCCCCCGGAGAGACGAGCTCCACCAAGCCAGCTCCCGCCAGCCCGCA 3004  
Db 2643 CATTAGGCCCCCGGAGAGACAGTACCCCAAGCCAGTTCCTCCGTGTCTCAACC 2702  
Qy 3005 CCGGAAACCTGACCGCGGAGAGAGCGCCCTCAAGCAGCTGTGACAGGCCCCCGGAGC 3064  
Db 2703 ACAGCAGACCTTACAGCCAGAGGAGTGAAGTGTCTCAGCAGTGTGAGAAAGTTCAGTGGC 2762  
Qy 3065 AAGAGCAGAGCCCGGACCCCGCGGACAAAGG-----G 3100  
Db 2763 AAGTCCCGAGAGCCAGTGTCTCTGCGAAGAAAGGACAGAAACCGGCACTTCTTTCG 2822  
Qy 3101 GCCTTTCAGCGGAGGCGCCAGAGCTGCTGTGGAGACCCCGCTTGTGACTTTCGAGCTG 3160  
Db 2823 GCTTTCACCACTGAGGGGCCAAAGCTAACGACTGAGCCCCCAACGCTGTGTACTGGGCTG 2882  
Qy 3161 CCTTCCCGTGTGCCCCCGGTGAGTATCAAGGCTTCCCGCAATGCCCGGAGCCCTCA 3220  
Db 2883 CCTTCCCGTGTGCCCCCGGTGAGTATCAAGGCTTCCCGCAATGCCCGTGTGAGCCCTCT 2942  
Qy 3221 GCCTTTCCTAAGCTCAACCTGATCCACCTGCGGCTGTGAGGCTGTGACAGTGTGCGG 3280  
Db 2943 GCTTTCCTCAACACCCCGGTGACCCGCTGCTGTGAGGCTGTCAAGATGTGCGCG 3002  
Qy 3281 CCGTTCCTGCGCGCGCCACCAACCATCTCCAACCGGCTCCCTCATCTCTGTCCAG 3340  
Db 3003 CCGTTCCTGCGCG---TCCCCCATCTCTAACCCCCCAACCCCTCATCTCTGTCCAG 3059  
Qy 3341 CACCCAGCGCTCTCGAGAGGCAAAATAGTGCATCTCC---CAAGAAATGTGTGTCAAG 3397  
Db 3060 CATCCCGGCTTACTTGAAGGAGAGTGTGTGCATCTCCAGAGGGAATGTCAGTCAAG 3119  
Qy 3398 CTCAAGCTCCCTGATCTCAAGCATGTCCAAAGGCCCCGGGTGTGTCTGTCACTATGGGCTG 3457  
Db 3120 CTTCGTGTGCTCACTCAAGCATGTCCAAAGGCCCCCATGTGTGTCTCACTATGGGCTG 3179  
Qy 3458 CCGCTGCGCATGAGCCCAAAAAGTGTGACACTTCAAGCGAGTGAAGCAGAGCAGCTG 3517  
Db 3180 CCGCTTGCCTGTGACCTTAAGAGCTG----- 3206  
Qy 3518 TCCCAAGGAGGCGAGCTGTGGCCACCGAGAGCCTGTGGGTGTCCACAGCCAGAGGCG 3577  
Db 3207 ----- 3206  
Qy 3578 TCCGTGTGAGAGGAGACAGCTTGTGGCTCAGTTCCGGGCGGAAGCATCAACAAAGCATT 3637  
Db 3207 -----GAGACAGACTGTGGCTTCGCGCACAGTGTGAAGCATCAACAAAGGCTC 3254  
Qy 3638 CCGAGACAGCGGAGGCTCTGAGACAGGCGCATCATACGCGGCTCATACCCAGCG 3697  
Db 3255 CCGAGTACCGGG-----CTGACAGCGGCCCGGCTCAAGAGGCTTATCAACCCAGCG 3308  
Qy 3698 AGCGCAGCTGACGTCTGTATCAAGAGGACCATCAACAGATCATGTGCGAGAGCAGCG 3757  
Db 3309 AGCGCGCAGACGTCTCTACAAAGGATACCATCAAGAGATCGTGTGTGAGAGACAGCCA 3368

Qy 3758 AGTGTCTTGAACCGCGGCGGAGAGACAGCTGTGCCAAAGGCGCATGTCTATCAAGAGC 3817  
Db 3369 AGTGCCTTGAACCGGAGACGAGAGACACCTGTGCCAAAGGCGCATGTCTATGAGAGC 3428  
Qy 3818 AAGAAAGGCCAGTCTTGTCTTATGAGGTTGACATGTCTGTGACCCAGTGTCCAAAGAG 3877  
Db 3429 AAGAAAGGCCAGTCTTATCTTATGAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3488  
Qy 3878 GAGGCGAAGAGACACTAGAGACCCCGCATGAGAGCGCGCGCGCCCAAGCGGCACTATGAC 3937  
Db 3489 GATGAAAGAGACAGCTGTGGGCGACCCCATGAGCTGCGCGCTTAAAGCATCTATGAC 3548  
Qy 3938 ATGATGAGGAGCGCGTGTGGGAGAGCATCTCTCAGCCAGCATGAGTCTCATGTGAGC 3997  
Db 3549 ATGATGAGGAGCGCGTGTGAGAGAGCTGTACCTCAACGACAGATAGAGGAGCTATGGC 3608  
Qy 3998 GTGTCCATCTCCGCGGAGGACAGAGCCCGCACCTCAAAAGAGCAGCACACATCCGC 4057  
Db 3609 CGGCGCATCTCCGTGAGCAGC-----ACAGCCCGCATCAAGGAGCAGCATCATCCGA 3662  
Qy 4058 GGTTCATCAACAAAGGATCCCTCGGTCTCATGTGAGGACAGAGAGATTAAGTGTGT 4117  
Db 3663 GGTTCATCAACGAAAGCATCTCCAGGATCTATGTGAGAGGCGAGAGAGTACTTGTGAG 3722  
Qy 4118 CGGAGGCGCAAGCTCTTAAAGGAGGAGGAGCGGCTCGCGCCCGGCGGCTCAAGGAGC 4177  
Db 3723 CGGAGGCGCAAGCTCTTGAAGCGAAGAGAGACACACACCCCGACACCTCGGAGC 3782  
Qy 4178 CTGACCGAGGCTTACAAAGCAGAGCC-----CTGGGCGCCCTGAAGCTGAAGCG 4228  
Db 3783 CTGACTGAGACTTACAAAGCCCGGCGCTGTGAGCCCTGTGGGTCTCTGAAGCTGAAGCG 3842  
Qy 4229 GCCATGAGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4288  
Db 3843 ACTCAACAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3902  
Qy 4289 CGGAGGAGCTGTGCGGACAGCGCCGAGCTGCGGCTGTGCGCGGCGGCTCAAGAGGAGC 4348  
Db 3903 AGAGAGAGCTGTGCGGACAGCTGTGAGCTACCTGTGAGACACAGGCTGTGAAGAGGAGT 3962  
Qy 4349 TCCATCAACGAGGAGCACCGCGCTCAAGTACAGACCGGCGGCTCACACTGTGCTCAAA 4408  
Db 3963 TCCATCAACGAGGAGCACCGCATCAAGTACAGCTGTGGGACACCTGTCACTGTGACCAAG 4022  
Qy 4409 AAGCAGACGATGCTCTCTCATGTGAGCGCCCGGCTGTGACGTTTCCACCCGTGTGACCC 4468  
Db 4023 AAACACGACGTTGCTTCATCATGTGAGCGCCCGGCGGCTTTCCTGTGCTGTGACCCG 4082  
Qy 4469 CTGATGTGATGTGCGGACGCGCGGCGGACCTGGAAGGTGTGCTTACAGAGAGGCTGAG 4528  
Db 4083 CTGACATTAATGTGTGACGCGCGGCGACCTGAGGTGTCTGTATGAAGAGTCTAAG 4142  
Qy 4529 AGCGGCGCAGAGACCGGACAGACTCGGAGGAGCTCAATTGTGCGCGGCGCGCGGCTCATT 4588  
Db 4143 AGCGGCTCAGAGACAGAGTGTGTGAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4282  
Qy 4589 GTGTCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4648  
Db 4203 GTGTCTGAACTGTGGCAAGCGACGCGCAAGCCCACTGACTTACGAAGACCAAGGAGCACCC 4262  
Qy 4649 TTTGTGCGGCGACCTCCACAGAGTGTGCGCGGTGTACCATGTGAGGAGCCAGCGCGGCTG 4708  
Db 4263 TTACACAGTACCTGTGACGCTGTGCTCTCTGTGTACCAAGAGGAGCCAGCGACGCTT 4322  
Qy 4709 CAGAGGCGAGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4768  
Db 4323 CAGAGAGGAGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4382  
Qy 4769 GTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4828  
Db 4383 CGGAGAGTGTGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4442

QY	4829	CCCTAATGAGCACTGCTCTCGGGGGGTGAGAGTGGCGTGGACCTGTATGCGACGCAATCCCC	4888
Db	4443	CCCTATATGAGCACTTGTCTCCGGGGGGGTGATCTGGTGTGACCTGTATCCGTGTGCATATCCA	4502
QY	4889	CTGGACCTTGCAGCCCCCACTCCATATACCCGCGGACATCCCTCTGGA---CGAGCGCGTCC	4945
Db	4503	TTGGCCCTTTGACCCCACTCCATATACCCGAGGATCCCTCTGGAAGAGACGCGGACGCC	4562
QY	4946	TACTACTGTCCCGAGCACTTGGCCCCCAACCCCACTTACCCTGACCTTGTACCACTTAC	5005
Db	4563	TACTACTGTCCCGGACCTTGGCCCCCAGCCCCCACTTACCACACTGTATCCCACTTAC	4622
QY	5006	CTCATTCGGCGGCTATCCCCGACACGGCGCGCGGTGGAGAAACCGGAGACATCATTAAGAC	5065
Db	4623	CTCATTCGGCGGCTATCCCTTGACACGGCGCGCGCTTGAAGAACCGGAGACATCATTAAGAC	4682
QY	5066	TACATCACTGTGAGAGATGACACCAACAACCGGCAACCGGCATATGGCCAGCGAGCTGAT	5125
Db	4683	TACATCACTGTGAGAGATGACACCAACAACCGCTGTCTCCGCAATGGCCAGCGAGCTGAC	4742
QY	5126	ATGCTGAGGGGCTCTCGCCCGCGAGTCTCGTGGACATCACTACGCTGCGGGTCCC	5185
Db	4743	ATGCTGAGGGGCTGTGTCAACCGCGAGAGTCTCGTGGCCCTCAATTAATGCGGTGGCCCA	4802
QY	5186	CGAGGCACTATGACACTGTCCCAAGTGACACACTCGCTGTGTCTGTGCCCCGACACCA	5245
Db	4803	AGAGGCAATTATCGACTGTGTCCAAAGTGACACACTCGCCGTGTGTGTGCCAACACGCCA	4862
QY	5246	GGCAACCCAGCAACGGCATGAGACCGACTTGTCTATCACTGCCCAACGGGCCCAACCTTC	5305
Db	4863	GGCAACCCCTGCGCACCGCATGACCGCTTGTGTACTCTCCCACTGGGCCCAACCTTTC	4922
QY	5306	AGCAGCCGCGCACAGACAGTCTCCCACTCTCCACAGAGGTGTCCACAACCTTGACAAAACA	5365
Db	4923	AGCAGCCGCGCACAGATGACTCAACCGGTGTCCCAAGAGGCCCACTCAACTATGATTAACA	4982
QY	5366	ACCAACCACTCTGTCTGTCCAGCGGGAGCGAGACCGGATCGAGAGCGGGACCCGGAGTCTGG	5425
Db	4983	ACTCCCACTATTTCATCGAGACGG-----GAAACGGGAACGTGAACGG	5024
QY	5426	GAGCGGGGAAAGTCCATCTCTACGTTCCACACAGACGGTGTGAGCAACGACCATCTGGAGA	5485
Db	5025	GAAAGAGCAAGTCCATCTCACTGTCTACACTACATAGTGTGACATGCCATCTTGGAGA	5084
QY	5486	CCTGGTACAGAGCAGAGCAGCGGACGACCGCGCAGCAACCGCGGGGTGTGGGGCAGACGC	5545
Db	5085	CCTGGTACGAGAGCAGAGCAGCGGGCTGTGGGGCAGACACCG-----	5125
QY	5546	AGCGCGCCCGCTCTCCACTCTCCATGTCCACCAAGACCTGTGCCCATCTCCCTGGACCCAG	5605
Db	5126	-----CCCGGCTCTCCACACCCACAGACACTGTGCCCATCTCCCCCGGACCCAG	5174
QY	5606	GATGCGCTTCCAGCAGAGACCCAGTGTGTTTACAACAAGGACATGAAGGGTATATATCAC	5665
Db	5175	GACGCTTTGACAGAGAGGCCACAGTGTGTGACACAAGACATATGAAGGGGTGTGTATCC	5234
QY	5666	GCTGTGAGAGCCAGACAAGCCACGGTCTCTG-----AGTTCACCTTCACTCTCTCA	5716
Db	5235	TTCGTGGAACCCGAGCAGCGCACGGTCTCTGAAGTGGGCAAGTTCACCTTCACTCTTTCG	5294
QY	5717	CCCGTTGCCAGCTGCAACATTCGCCACTGTGCACCACTGTGCCCACTTGGGCGGACCTCTC	5776
Db	5295	CCTGTCCGCCAGCTGCACATTTCCCACTGTGCACCACTGTGCCCACTTGTGTGGACCTTT	5354
QY	5777	GATGGGGGTCTAACCTTACCTCATTGAAGCCGCTTGTGTGCCCCAAGAGAGCCCCCGGGTC	5836
Db	5355	GAAAGGGGTCTAACCTTACCTCATTGAAGCCGCTCTGTTAACCAAGAGAACCTCTGTGGGTC	5414
QY	5837	GCCCGGACAGAGCGGCCCCCGAGACACAACCGGCACTGTCTTCTGTGCAAGACCCCCAGCC	5896
Db	5415	GCCCGGACAGAGCGGCCCCCGGGTGAAGCTGTGCATGTCTTTTCAACAAACCCCGGGCC	5474
QY	5897	CGCTCCGGGCTGAAGCCCGCTCTCTCCCAACAGCAAGGCTCGAAGCCCGGCGCTTATGT	5956

Db	5475	CG-----GGAGCCCGGCTCCTTACCCAGAAAGGCTCCGAGGCCCGATCCCTTAGCA	5525
QY	5957	CTCTCTGTCTTGGCCACGCGCAACATGCCCCGCAACCTTGGGAAAGACTTCGACCTTAC	6016
Db	5526	CCCCCAGCTTCAGGCCACAGGCGCATGCCCCGCAACCCAGCAAGAAACTTTCGACCCAC	5585
QY	6017	CAGCCGAGCCCGGACCCCGCGGCGCCACTGCGCTCCGCGCTCGGACCCGCAACCGGAAAG	6076
Db	5586	CATGCCAGTCCGAGCCCGCGGGG---CCACCTCGGCTCTAGATTCTGCACCGAGAAAG	5642
QY	6077	ACTCAAGTAAACCCCTTTTCCATTCAGAGAACTGGAACTCGTCTCTGCGGTTACAC---	6133
Db	5643	ACTCAAGTAAACCCCTTTTCCATTCAGAGAACTGGAACTCGGTTCTCTGCGGTTACACAGT	5702
QY	6134	GGCAGAGCTACAGACCCCGAAGGGGTGAGCCCGTCAGCCCTGTGAGCTCAACCACTCTG	6193
Db	5703	GGAGCTGGCTACAGCCCGGATGGGGTGGAGCCCATCAGCCCGGTGAGCTCCCCAGCCTTG	5762
QY	6194	ACCCACGACAAAGGGGCTTCCCAAGCACTTGAAAGAGTTCACAAAGAACCAACCTGGAGGG	6253
Db	5763	ACCCACGACAAAGGGGCTTCTCCAACTCTTGAAAGCTTACAGAAAGAACCACTTGAAAGGG	5822
QY	6254	GAGCTGGGGCCCAAGAGCCAGGCCCCGTGAAAGCTTGGCGGGAGGCGCGCCACTTCCA	6313
Db	5823	GAGCTGGGGCCCAAGAGCCAGGCCCCCATAGAGCTCAGCCGGAGGCGTGCCTCATCTCCA	5882
QY	6314	CACCTGGCGGCGCTGCTGAGAGCCAGCCCTCGTCCAGCCCGCTGTCCAGACCGCCCCA	6373
Db	5883	CATCTGGCGGCTCATGCCCCGAGAGCCACCTCATTCAGGCCCATCTCTCCAGATGCCCCCA	5942
QY	6374	GGGGTCAAAGGTCAACAGCGGGTGTATCCTTGGCCAGACAATCAGTGAAGTATATACA	6433
Db	5943	GGGATCAAAGGTCAACAGAGGGTGTATCCTTGGCTCAGACAATCAGTGAAGTATTTAGC	6002
QY	6434	CAGAGCTAACACCGGGCAACACCCACAGAGCTCAGGCCAACCCCTGCGCCGCCCTCTTAC	6493
Db	6003	CAGAGCTAACACCGGGCAACACCCAGAGCTCAGTGGCCCCCTTCCCGCCCTCTCTAC	6062
QY	6494	TCTTTCCTTGGGGCAGCTGCCCCGTCTGTGACCTTCGCGCGCCAGCCAGTGACTTTCAC	6553
Db	6063	TCTTTCCTTGGGGCAGCTGCCCCGTCTGTGATCTTTCGCGCGCCAGCCAGTGACTTTCAC	6122
QY	6554	CTCCCCGCCCCGACATGTGTGCCCGGCGCTGTGGTCTTCCCCCAACAGGAAAGGGGGCAG	6613
Db	6123	CTCCCCGCCCCGACATGTGCACTCCAGCCCGGGGAACTCCCCCAAGTGAAGGGGGCAAA	6182
QY	6614	AGGCTCCAGAGCCAAACAGACGTCGAGTCTTGGGTGTGGTGAAGACGGAATTGAAGCT	6673
Db	6183	AGGCTCCAGAGCCAGCAAAACATCGGTCTTGGGCGAGCGAGAGATGCCATTGAAGCT	6242
QY	6674	GTGTGCCACCGGAGGGGATGACCGAGGCACTCCCGAGTGTGTATCCGCTG	6733
Db	6243	GTGTGCCACCGGAGGGGATGATCTGAGCGAGGACATGTGTGAGACACTCGCTACCACTG	6302
QY	6734	CTGTATCCGGGATGCGGAAACAGACGAGCCACGACGATGGGCTCTCAAGTCTTCAGGCAAC	6793
Db	6303	CTGTATCGAGACGGGGGAAACAGGCGAGGCG---CAGATGTGGTCTTAGGTCTTCAGGCAAC	6359
QY	6794	ACCGAGCAGCCGCGAGCTTTCTTACGAAAGCTGACCGAGACCACTCCGCAATGCTCAAG	6853
Db	6360	ACCGAGCAGCCGCGCACTTTCTTACGTAAGCTGACCGAGACCACTCCGCAATGCTCAAG	6419
QY	6854	TCCAGAAAGCAGAGATTCACAAAGAGCTGAACCCACCAACCGGAAATGAGCTGAATAC	6913
Db	6420	TCCAGAAAGCAGAGATTCACAAAGAACTCAACCCACCAACCGGAAAGAGCCAGAAATAC	6479
QY	6914	AATATCAGCCAGCTTGAGCGGAGATTTCAATAATGCCCGCATACCGGAAACAGGCTTT	6973
Db	6480	AATATGCGCAGCTTGAGCGGAGATTTCAATAATGCCCGCATACCTGAGACAGGCTTT	6539
QY	6974	ATGACCTATAGAGCCAGGCGGTGACGAACATGTCAGACACCAATGGGGCTGGAGGCC	7033



Db 332 CATGTGGGATTCACACAGCCTGTGTGGCAGACGTGGAGGGCCACTGAGCCCGCTACCC 391  
Qy 61 GCCCCACAGCCTTTCTCTACCCAGTGCAGATCGCCGAGCAGCAGAGGTGGGCTCT 120  
Db 392 GCCCCACAGCCTTTCTCTACCCAGTGCAGATCGCCGAGAGCAGCAGAGGTGGGCTCT 451  
Qy 121 GAGATTCAGACACACTCCCGACTATGCTCCCACTGTGCGCGGGCTTCATATCCA 180  
Db 452 GAGATTCAGACACACTCCCGACTATGCTCCCACTGTGCGCGGGCTTCATATCCA 511  
Qy 181 GCCCGAGGGGGGAGGGCCCTCCCTGCTGTGTGAGTTCAGACCCGGGAATGAAAGGTCCCA 240  
Db 512 GCCCGAGGGGGGAGGGCCCTCCCTGCTGTGTGAGTTCAGACCCGGGAATGAAAGGTCCCA 571  
Qy 241 GAGAGTTCACCTGCGGGCAGAGTCCCACTCATACCTGCCGAGCTGGGGAATCAGAGAT 300  
Db 572 GAGAGTTCACCTGCGGGCAGAGTCCCACTCATACCTGCCGAGCTGGGGAATCAGAGAT 631  
Qy 301 GAGAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGTGCTGACCCCTGTGCGAAC 360  
Db 632 GAGAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGTGCTGACCCCTGTGCGAAC 691  
Qy 361 GTTCAACCCCTGTGTGCGACCGGGCAGCCTGCGGGATCTGAAAGCTCACCAAGAACCTTAG 420  
Db 692 GTTCAACCCCTGTGTGCGACCGGGCAGCCTGCGGGATCTGAAAGCTCACCAAGAACCTTAG 751  
Qy 421 CCTGACGGGCAAGCTGGAACCGGATGTTCTCCCGCAGCGCCGCACTGACCTGAGCT 480  
Db 752 CCTGACGGGCAAGCTGGAACCGGATGTTCTCCCGCAGCGCCGCACTGACCTGAGCT 811  
Qy 481 GAGAGTGTGCTCCGCAACCGGCTGTCCAGAGAGAGCTGATCCAGAACATGACCGGCTGGA 540  
Db 812 GAGAGTGTGCTCCGCAACCGGCTGTCCAGAGAGAGCTGATCCAGAACATGACCGGCTGGA 871  
Qy 541 CCGAAGATTCACAGATGAGAGACAGACAGATCTCTAAGCTGGAAGAGAGAGACAGACAGT 600  
Db 872 CCGAAGATTCACAGATGAGAGACAGACAGATCTCTAAGCTGGAAGAGAGAGACAGACAGT 931  
Qy 601 GAGAGAGAGAGCTGCGCAAGCGCCGCGAGCTGAGAAAGCCGTTGTACCGCGCCCATGGA 660  
Db 932 GAGAGAGAGAGCTGCGCAAGCGCCGCGAGCTGAGAAAGCCGTTGTACCGCGCCCATGGA 991  
Qy 661 GTTGAAGCACCAGCCTGTGTGCAAGATCATCTTACAGAGAAACCGGAAGAGGCTGGAAG 720  
Db 992 GTTGAAGCACCAGCCTGTGTGCAAGATCATCTTACAGAGAAACCGGAAGAGGCTGGAAG 1051  
Qy 721 TGCAATGAGGATTTGTGAAAGGCTGTGGGGCCCGAGGTGAGCTGCCGCTGTACAAACGACC 780  
Db 1052 TGCAATGAGGATTTGTGAAAGGCTGTGGGGCCCGAGGTGAGCTGCCGCTGTACAAACGACC 1111  
Qy 781 CTCGCAACACCCGCGCAGTATCATGAGAACATCAAAATTTAAACAGAGGATCGGAAAGACT 840  
Db 1112 CTCGCAACACCCGCGCAGTATCATGAGAACATCAAAATTTAAACAGAGGATCGGAAAGACT 1171  
Qy 841 AATCTTGTATCTTCAAGAGAGGAAATCAGGCTCGAAACATGGAAGCAGAGATTCTGCCA 900  
Db 1172 AATCTTGTATCTTCAAGAGAGGAAATCAGGCTCGAAACATGGAAGCAGAGATTCTGCCA 1231  
Qy 901 GCGCTATATACCAAGCTCATGAGAGGCTTGTGAAAATAAGGTGAGCCCATGTGAAAACAACC 960  
Db 1232 GCGCTATATACCAAGCTCATGAGAGGCTTGTGAAAATAAGGTGAGCCCATGTGAAAACAACC 1291  
Qy 961 GCGCGCGCGCGCCAGAGAGAGAGAGAGTGGCGCAGTACTAGAAAAGAGTTCTCTGAGAT 1020  
Db 1292 CCGCGCGCGCGCGCCAGAGAGAGAGAGAGTGGCGCAGTACTAGAAAAGAGTTCTCTGAGAT 1351  
Qy 1021 CCGCAAGCAGCGCGCAGCTGCAAGAGCGCATGCAAGCAGAGGTTGGCGCAACCGGCGCATG 1080  
Db 1352 CCGCAAGCAGCGCGCAGCTGCAAGAGCGCATGCAAGCAGAGGTTGGCGCAACCGGCGCATG 1411  
Qy 1081 GCTGTCCATGTCGCGCGCGCGCGAGAGAGCAGAGGTGTCAGAGATCATGAGGCTCTGTC 1140  
Db 1412 GCTGTCCATGTCGCGCGCGCGCGAGAGCAGAGGTGTCAGAGATCATGAGGCTCTGTC 1471

Qy 1141 AGAGCAGAGAACTTGAAGAGCAGATGCGCAGCTGGCCGTTGATCCGCCATGCTGTA 1200  
Db 1472 AGAGCAGAGAACTTGAAGAGCAGATGCGCAGCTGGCCGTTGATCCGCCATGCTGTA 1531  
Qy 1201 GAGCCTTACCAAGCAGCGCATCAAGTTTATCATCACTTAAAGGGCTTATGGCCGACCCAT 1260  
Db 1532 GAGCCTTACCAAGCAGCGCATCAAGTTTATCATCACTTAAAGGGCTTATGGCCGACCCAT 1591  
Qy 1261 GAAGGTATCAAAAGACCGCCAGGTATGAAATGATGAGTGAAGAGAGAGACCTT 1320  
Db 1592 GAAGGTATCAAAAGACCGCCAGGTATGAAATGATGAGTGAAGAGAGAGACCTT 1651  
Qy 1321 CCGGAGAAAGTTTATGACAGATCCCAAGAACTTTGGCCTGATCCGATCTTGGAGAG 1380  
Db 1652 CCGGAGAAAGTTTATGACAGATCCCAAGAACTTTGGCCTGATCCGATCTTGGAGAG 1711  
Qy 1381 GAAACAGTGGCTGAGTGCCTCTTATTACTTACTGTAAGAAAGAAATGAAATCTTAA 1440  
Db 1712 GAAACAGTGGCTGAGTGCCTCTTATTACTTACTGTAAGAAAGAAATGAAATCTTAA 1771  
Qy 1441 GAGCCTGTGAGACGAGCTATTCGGCGCGCGGCAAGAGCCAGCAGCAACACAGACA 1500  
Db 1772 GAGCCTGTGAGACGAGCTATTCGGCGCGCGGCAAGAGCCAGCAGCAACACAGACA 1831  
Qy 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
Db 1832 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1891  
Qy 1561 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
Db 1892 AGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1951  
Qy 1621 CGACAGAGAAAGCTCTCAAGAGAGAAAGACAGACACTCAAGGGAGAGCAACAGAGA 1680  
Db 1952 CGACAGAGAAAGCTCTCAAGAGAGAAAGACAGACACTCAAGGGAGAGCAACAGAGA 2011  
Qy 1681 GAAAGAGCTGTGGCTTCCAAAGGCGGCAAACTGTCAAACAGCCAGGAGAGCGCAAGG 1740  
Db 2012 GAAAGAGCTGTGGCTTCCAAAGGCGGCAAACTGTCAAACAGCCAGGAGAGCGCAAGG 2071  
Qy 1741 CCGCATCACCCGCTCAATGAGCTTAATGAGGCCAACAGAGAGAGGCTATCACCCCAACA 1800  
Db 2072 CCGCATCACCCGCTCAATGAGCTTAATGAGGCCAACAGAGAGAGGCTATCACCCCAACA 2131  
Qy 1801 GAGCGCCGAGGCTGCTCATGAGCTGAATGAGAGTTCTGCTGAGCAGAGAGAGAAAT 1860  
Db 2132 GAGCGCCGAGGCTGCTCATGAGCTGAATGAGAGTTCTGCTGAGCAGAGAGAGAAAT 2191  
Qy 1861 GGAACAGCCAAAGAAAGTCTCTGAGAACAGCGCCGCAACTGTGGCCATTCGCCGAGAT 1920  
Db 2192 GGAACAGCCAAAGAAAGTCTCTGAGAACAGCGCCGCAACTGTGGCCATTCGCCGAGAT 2251  
Qy 1921 GGTGGGCTTCAAGACTGTGTGCAAGTGTAAAGACTTCACTTCAACTCAAGAAAGAGCA 1980  
Db 2252 GGTGGGCTTCAAGACTGTGTGCAAGTGTAAAGACTTCACTTCAACTCAAGAAAGAGCA 2311  
Qy 1981 GAACTCGATGAGATCTTGAAGCAGCAGCAAGCTGAAGATGAGAGAGAGAGAGAGCGCG 2040  
Db 2312 GAACTCGATGAGATCTTGAAGCAGCAGCAAGCTGAAGATGAGAGAGAGAGAGAGCGCG 2371  
Qy 2041 GAGAGAAAGAAAGAAAGCGCGCGCGCGCGCAGCAGAGAGGCTGCAATTCGCGCCGTGGT 2100  
Db 2372 GAGAGAAAGAAAGAAAGCGCGCGCGCGCGCGCAGCAGAGAGGCTGCAATTCGCGCCGTGGT 2431  
Qy 2101 GAGAGATGAGAGAGATGAGAGGCTGCGGCTGAGCGGAAATGAGAGAGAGATGAGTGAAGA 2160  
Db 2432 GAGAGATGAGAGAGATGAGAGGCTGCGGCTGAGCGGAAATGAGAGAGAGATGAGTGAAGA 2491  
Qy 2161 GAGCTGAAGCTTTACATGCTCTGGGAATGAGGTGCCAGAGGGGAATCAGTGGCCAGC 2220  
Db 2492 GAGCTGAAGCTTTACATGCTCTGGGAATGAGGTGCCAGAGGGGAATCAGTGGCCAGC 2551







COJ22208 LOCUS COJ22208 6339 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 8142 from Patent WO02068579.  
ACCESSION COJ22208  
VERSION COJ22208.1 GI:42283065  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE  
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
Kits, such as nucleic acid arrays, comprising a majority of  
humanecons or transcripts, for detecting expression and other uses  
hereof  
JOURNAL Patent: WO 02068579-A 8142 06-SBP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source  
1..6339  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 46.3%; Score 3963.8; DB 6; Length 6339;  
Best Local Similarity 75.6%; Pred. No. 0;  
Matches 5911; Conservative 0; Mismatches 167; Indels 1737; Gaps 11;  
2 ATGTGGGGCTCCACACAGCTTGTGGACAGAGTGAAGGGCCACTGAGCCCCGCTAACCG 61  
1 AAGTGGGATCCACACAGCTGTGGACAGAGTGAAGGGCCACTGAGCCCCGCTAACCG 60  
62 CCCACAGGCTTCTCTACCAAGTGCAGATGCGCCGACGACACAGGACGTGGGGCTCTG 121  
61 CCCACAGGCTTCTCTACCAAGTGCAGATGCGCCGACGACACAGGACGTGGGGCTCTG 120  
122 GAGTACAGACACACATCCCGCAGACTATGCTCCACCTGTGCGCGGGCTCATCATCCAG 181  
121 GAGTACAGACACACATCCCGCAGACTATGCTCCACCTGTGCGCGGGCTCATCATCCAG 180  
182 CCCACAGGCGGAGGCGCTCTGCTGTGTGAGTTCCAGCCCGGAAATGAAGGCTCCAG 241  
181 CCCACAGGCGGAGGCGCTCTGCTGTGTGAGTTCCAGCCCGGAAATGAAGGCTCCAG 240  
242 GAGCTCCACCTGCGGCGAGAGTCCCATGATACCTGCGCGAGCTGGGGAATCAAGATG 301  
241 GAGCTCCACCTGCGGCGAGAGTCCCATGATACCTGCGCGAGCTGGGGAATCAAGATG 300  
302 GAGTTCATTGAAGAAGGAGCGCTCGGCTAAGAGCTGCTGACCCCTGCTGCGACCG 361  
301 GAGTTCATTGAAGAAGGAGCGCTCGGCTAAGAGCTGCTGACCCCTGCTGCGACCG 360  
362 TCACCCCTGCTGCGCACAAGGCGCAGCTGCGGATCTGAAGACCTCAACAGAGCGGTAGC 421  
361 TCACCCCTGCTGCGCACAAGGCGCAGCTGCGGATCTGAAGACCTCAACAGAGCGGTAGC 420  
422 CTGAGGGGCAAGCTGGAACCGGTGTCTCCCCCAAGCCCCCGACACATGACCTTAGCTG 481  
421 CTGAGGGGCAAGCTGGAACCGGTGTCTCCCCCAAGCCCCCGACACATGACCTTAGCTG 480  
482 GAGCTGTGCGCCACAGGCTGTCCAAAGAGAGTGTACAGAACTGAGACCGCTGAC 541  
481 GAGCTGTGCGCCACAGGCTGTCCAAAGAGAGTGTACAGAACTGAGACCGCTGAC 540  
542 CGAGAGATCAACATGATAGAGAGAGATCTTAAAGCTGAAGAAGAGAGAGAGAGCTG 601  
541 CGAGAGATCAACATGATAGAGAGAGATCTTAAAGCTGAAGAAGAGAGAGAGAGCTG 600  
602 GAGGAGAGAGCTGCGCAAGCGCGCGAGCTTGAAGAGCCCGGTGACCCGCGCATTCAG 661  
601 GAGGAGAGAGCTGCGCAAGCGCGCGAGCTTGAAGAGCCCGGTGACCCGCGCATTCAG 660

QY 662 TCGAAGCACCCGAGCTGTGTGATCATCTAACGAGAACCCGGAAGAGGCTGAGCT 721  
DB 661 TCGAAGCACCCGAGCTGTGTGATCATCTAACGAGAACCCGGAAGAGGCTGAGCT 720  
QY 722 GCACATCGAATCTGGAAGGCTGGGGCCCGCAGGTGAGCTGCGCTGTACCAACAGCC 781  
DB 721 GCACATCGAATCTGGAAGGCTGGGGCCCGCAGGTGAGCTGCGCTGTACCAACAGCC 780  
QY 782 TCCGACACCCGAGATCATGAGAACATCAAAATAAACAGCGCATGCGGAAGACTA 841  
DB 781 TCCGACACCCGAGATCATGAGAACATCAAAATAAACAGCGCATGCGGAAGACTA 840  
QY 842 ATCTGTACTTCAAGAGAGAAATCAGCTGCGAAACATGGAACCAAGTTCTGCCAG 901  
DB 841 ATCTGTACTTCAAGAGAGAAATCAGCTGCGAAACATGGAACCAAGTTCTGCCAG 900  
QY 902 CGCTATGACCAAGCTATGAGAGGCTTGGAAAAAAGGTGAGCGCATGAAAAACAACCG 961  
DB 901 CGCTATGACCAAGCTATGAGAGGCTTGGAAAAAAGGTGAGCGCATGAAAAACAACCG 960  
QY 962 CGCCGCGGCGCCAGAGAGAGCAAGGTGCGAGATCTACGAAAGCAAGTCCCTGAGATC 1021  
DB 961 CGCGGCGGCGCCAGAGAGAGCAAGGTGCGAGATCTACGAAAGCAAGTCCCTGAGATC 1020  
QY 1022 CGCAAGCAGCGGAGCTGAGAGAGCGCATTCAGAGACAGGCTGGCCAGCGGGGCGTGG 1081  
DB 1021 CGCAAGCAGCGGAGCTGAGAGAGCGCATTCAGAGAGCGCATTCAGAGAGCGGCGTGG 1077  
QY 1082 CTGTCAATGTGGGCGCGCGCAGAGAGAGAGGATGATGATGATGATGATGATGATGAT 1141  
DB 1078 CTGTCAATGTGGGCGCGCGCAGAGAGAGAGGATGATGATGATGATGATGATGATGAT 1137  
QY 1142 GAGCAGAGAACTTGAAGAGAGATGCGCAGAGCTGCGCTGATCCGCGCATGCTGTAC 1201  
DB 1138 GAGCAGAGAACTTGAAGAGAGATGCGCAGAGCTGCGCTGATCCGCGCATGCTGTAC 1197  
QY 1202 GACGCTGACAGAGAGCGCATCAAGTTCAATCAACATGAAAGGCTTATGAGCGACCCATG 1261  
DB 1198 GACGCTGACAGAGAGCGCATCAAGTTCAATCAACATGAAAGGCTTATGAGCGACCCATG 1257  
QY 1262 AAGGTTCAAAGAACCGCAGAGTCAATGAAACATGAGAGAGAGAGAGAGAGAGAGAG 1321  
DB 1258 AAGGTTCAAAGAACCGCAGAGTCAATGAAACATGAGAGAGAGAGAGAGAGAGAGAG 1317  
QY 1322 CGGAGAAATTCATGACACATCCCAAGAACTTGGCTGATGATCATCTCTGAGAGAG 1381  
DB 1318 CGGAGAAATTCATGACACATCCCAAGAACTTGGCTGATGATCATCTCTGAGAGAG 1377  
QY 1382 AAGACAGTGGCTGAGTGGCTCTTATTAATTAATTAATTAATTAATTAATTAAG 1441  
DB 1378 AA----- 1379  
QY 1442 AGCTGTGAGAGAGAGCTATGCGGCGCGCAGAGACGACAGCAACACAGAGCAG 1501  
DB 1380 -----GATGAGTGGCTGCGCGCGCATACCTTTC 1409  
QY 1502 CAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1561  
DB 1410 GGTCTTCACTTCCGTGGGCACTGCTGAGACTCAG----- 1445  
QY 1562 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621  
DB 1446 ----- 1445  
QY 1622 GACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681  
DB 1446 -----GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1485  
QY 1682 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1741  
DB 1486 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1545  
QY 1742 CGCATCAACCGCTCAATGAGCTAATGAGGCAACAGAGAGAGAGAGAGAGAGAGAGAG 1801

Db 1546 CGCATCACCCGCTCAATGGCTAATGAGCCCAAGCGAGAGGCTATCAACCCCCAGAG 1605  
QY 1802 AGGCGGAGGTGGCCCTCCATGAGAGCTGTAATGAGATTCTGGCTGGAATAAGAAATG 1861  
Db 1606 AGGCGGAGGTGGCCCTCCATGAGAGCTGTAATGAGATTCTGGCTGGAATAAGAAATG 1665  
QY 1862 GAACAAGCCAAAGAAAGTCTCTCTGAAACAGGCGCAACTGTCTGGCCATCGCCGAGT 1921  
Db 1666 GAACAAGCCAAAGAAAGTCTCTCTGAAACAGGCGCAACTGTCTGGCCATCGCCGAGT 1725  
QY 1922 GTGGGCTCCAAAGCTGTGTGCAAGTGTAAAGCTTCTTCACTTCAATCAAGAAAGGAG 1981  
Db 1726 GTGGGCTCCAAAGCTGTGTGCAAGTGTAAAGCTTCTTCACTTCAATCAAGAAAGGAG 1785  
QY 1982 AACCTGATGATGATCTTTGAGAGAGCAACAAGCTGAAGATGGAAGAAGAAAGGAGAG 2041  
Db 1786 AACCTGATGATGATCTTTGAGAGAGCAACAAGCTGAAGATGGAAGAAGAAAGGAGAG 1845  
QY 2042 AGGAAGAAGAAAGAGCGCGCGGCGCGAGCGAGAGAGGCTGATTCCCGCCGCTGTGT 2101  
Db 1846 AGGAAGAAGAAAGAGCGCGCGGCGCGAGCGAGAGAGGCTGATTCCCGCCGCTGTGT 1905  
QY 2102 GAGATGAGAGAGATGAGAGCGGTGCGGCGTGAAGCGAAATGAGAGAGATGTGAGAGAG 2161  
Db 1906 GAGATGAGAGAGATGAGAGCGGTGCGGCGTGAAGCGAAATGAGAGAGATGTGAGAGAG 1965  
QY 2162 GCTGAAGCTTAACTGCTCTTGGGAATGAGGTGCCAGAGAGGAAATGCAATGGCCAGCC 2221  
Db 1966 GCTGAAGCTTAACTGCTCTTGGGAATGAGGTGCCAGAGAGGAAATGCAATGGCCAGCC 1993  
QY 2222 ACTGTCAACAACAGCTCAGACACGAGAGATCCCTCTCTCAACTGAGGCGGCAAG 2281  
Db 1994 ----- 1993  
QY 2282 GACACAGGCGAGATGAGGCCCAAGCCGCCAGCAACCTTGGCGCCGAGCGGCAACCCCA 2341  
Db 1994 ----- 1993  
QY 2342 GGGCCACCAACCCCAACGAGAGACATCCGGGCCCCCATTTGAGCCACCCCGGCTCT 2401  
Db 1994 ----- 1993  
QY 2402 GAAGCAACGAGAGCCCTTACGCCCCCAACAGCAACCCCATCGCTCTGTGACCTCTCT 2461  
Db 1994 ----- 1993  
QY 2462 GTGGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2521  
Db 1994 ----- 1993  
QY 2522 GAGGAGCAAGAGCCCCCGCGGCTGAGAGAGCTGGCAAGTGAACAAGGAAAGCCGAGAGAG 2581  
Db 1994 ----- 1993  
QY 2582 CCCGTCAAGAGAGAGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2641  
Db 1994 ----- 1993  
QY 2642 GAGGCGCTGAGAGCAACGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2701  
Db 1994 ----- 1993  
QY 2702 GCGAGAGCAACATGCGCAAGAGAGTGGGCGGCCCCCAGAGCAAGCGATCTCAAGTCTACC 2761  
Db 1994 ----- 2016  
QY 2762 TGCAATGACAGAGAGATGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2821  
Db 2017 TGCAATGACAGAGAGATGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2076  
QY 2822 AGGCGAGGCTCTCAACCCGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2881

Db 2077 AGGCGAGGCTCTCTCAACCCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2136  
QY 2882 CTGGAACCTGAAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2941  
Db 2137 CTGGAACCTGAAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2184  
QY 2942 GTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3001  
Db 2185 ----- 2184  
QY 3002 CCAACCGCAAACTGCAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3061  
Db 2185 ----- 2184  
QY 3062 GGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3121  
Db 2185 ----- 2191  
QY 3122 AGCTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3181  
Db 2192 AGGCTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2225  
QY 3182 GAGGTATCAAGGCTTCCCGCATGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3241  
Db 2226 ----- 2232  
QY 3242 GGTCAACCACTGGCGGCTTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3301  
Db 2233 GGTCAACCACTGGCGGCTTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2292  
QY 3302 ACCATCTCAACCGGCTCCCTCATCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3361  
Db 2293 ACCATCTCAACCGGCTCCCTCATCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2352  
QY 3362 CAATAGTGTGCATCTTCCAGAGAGATGTGCTCAAGTCTCAAGTCTCTCAAGAGCAT 3421  
Db 2353 CAATAGTGTGCATCTTCCAGAGAGATGTGCTCAAGTCTCTCAAGAGCAT 2373  
QY 3422 GCCAAGGCCCCGCTGGGCTTCTCAACATGGGCTGCCCCGATGAGAGAGAGAGAGAGAG 3481  
Db 2374 ----- 2373  
QY 3482 CTGGCAACCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3541  
Db 2374 ----- 2373  
QY 3542 CCGAGAGCTTGGAGGTGCCACAGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3601  
Db 2374 ----- 2373  
QY 3602 GGCTAGTTCGGAGCGGAAGCATCAACAAAGGATTCAGAGAGAGAGAGAGAGAGAGAGAG 3661  
Db 2374 ----- 2373  
QY 3662 AGGCGCATCATACCGCGGCTCATCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3721  
Db 2374 ----- 2400  
QY 3722 GGCACCATCAACAGATCATTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3781  
Db 2401 GGCACCATCAACAGATCATTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460  
QY 3782 GACAGCTTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3841  
Db 2461 GACAGCTTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
QY 3842 GAGGCTGAGATGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3901  
Db 2521 GAGGCTGAGATGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
QY 3902 CCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3961  
Db 2581 CCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640

QY 3962 GGCATCTCTCAGCCAGCATCGAAGGTCTCATGAGCCGTGCCATCCCGCCGAGCGACAC 4021  
DB 2641 GGCATCTCTCAGCGACGATCGAAGT----- 2667  
QY 4022 AGCCCCCACACCTCAAAAGACGACACATCCCGGGTCCATCACACAGGATCCCT 4081  
DB 2668 ----- 2667  
QY 4082 CGGTCTTACGTGAGGACACAGGAGCTACCTGCGTCCGAGGAGCCAAAGCTCTAAAGCGG 4141  
DB 2668 ----- 2667  
QY 4142 GAGGACAGCCTCCGCCCCCACCGGCTTCACCGGACCTTGACCGAGGCTTACAAAGACGACG 4201  
DB 2668 GATTACAGGAGGAGAGCTTTCATCTCTCGGTGCCCTGTGTGGCGGGCTTACAAAGACGACG 2727  
QY 4202 GCCCTGGGCCCCCTGAGAGCTGAAGCCGGGCTTGAAGGGCTGTGGTCCACGGTGAAGAG 4261  
DB 2728 GCCCTGGGCCCCCTGAGAGCTGAAGCCGGGCTTGAAGGGCTGTGGTCCACGGTGAAGAG 2787  
QY 4262 GCGGGCCGCTTCATCATGAGATCCCGGCGAGAGGCTGCGGACACGCGCGAGCTGCC 4321  
DB 2788 GCGGGCCGCTTCATCATGAGATCCCGGCGAGAGGCTGCGGACACGCGCGAGCTGCC 2847  
QY 4322 CTGGCCCCCGCGGCGCTCAAGAGAGGCTTCATCACGCAAGGACCCCGCTCAAGTACGAC 4381  
DB 2848 CTGGCCCCCGCGGCGCTCAAGAGAGGCTTCATCACGCAAGGACCCCGCTCAAGTACGAC 2907  
QY 4382 ACCGCGCGCTCCACCATCTGGCTTCCAAAAAGCACGATGCGTCCCTCATCGGCGACCCC 4441  
DB 2908 ACCGCGCGCTCCACCATCTGGCTTCCAAAAAGCACGATGCGTCCCTCATCGGCGACCCC 2967  
QY 4442 GCGCGGACGTTCCACCGGTGACCGCGCTGATGATGAGCGGACGCGCGGACCTGGA 4501  
DB 2968 GCGCGGACGTTCCACCGGTGACCGCGCTGATGATGAGCGGACGCGCGGACCTGGA 3027  
QY 4502 CGTGCCTGCTACGAGAGAGAGCTTGAAGACCGGACCGGACCGGACCGGACCTCGGAGG 4561  
DB 3028 CGTGCCTGCTACGAGAGAGAGCTTGAAGACCGGACCGGACCGGACCGGACCTCGGAGG 3087  
QY 4562 TTCATTGGCGCGGCGGCGCGGCTCATTTGCTGAGCTGGGTAAAGCCGCGGACGAGCCCC 4621  
DB 3088 TTCATTGGCGCGGCGGCGGCGGCTCATTTGCTGAGCTGGGTAAAGCCGCGGACGAGCCCC 3147  
QY 4622 CTGACCTTATGAGGACACAGGGGACCTTTGCGGACCTTCCAGAGGTTGCGCGGTG 4681  
DB 3148 CTGACCTTATGAGGACACAGGGGACCTTTGCGGACCTTCCAGAGGTTGCGCGGTG 3207  
QY 4682 ACCATGCGGAGCCCAACGCGCGGCTGACGAGAGGAGCTTTGCTCCAGAGGACATCC 4741  
DB 3208 ACCATGCGGAGCCCAACGCGCGGCTGACGAGAGGAGCTTTGCTCCAGAGGACATCC 3267  
QY 4742 CAGGACCGGAAAGCTGACGTGACGCGCTGTGAGATGCGCAAGTCCCCGACAGACCTTG 4801  
DB 3268 CAGGACCGGAAAGCTGACGTGACGCGCTGTGAGATGCGCAAGTCCCCGACAGACCTTG 3327  
QY 4802 CCGGAGCACCAACCAACCACTTTCGACCTTATGAGCACCTGCTTTCGAGGGCGTGAAGTGGC 4861  
DB 3328 CCGGAGCACCAACCAACCACTTTCGACCTTATGAGCACCTGCTTTCGAGGGCGTGAAGTGGC 3387  
QY 4862 GTGAGACTGTATGAGGACCAATCCCGCTGAGGCTTTCGACCCCACTTCAATCCCGGAGC 4921  
DB 3388 GTGAGACTGTATGAGGACCAATCCCGCTGAGGCTTTCGACCCCACTTCAATCCCGGAGC 3447  
QY 4922 ATTCCTCTGGAAGCAGCGCGCTGCTTATCTGCCCCGACACTTGGCCCCCAATCCCAAC 4981  
DB 3448 ATTCCTCTGGAAGCAGCGCGCTGCTTATCTGCCCCGACACTTGGCCCCCAATCCCAAC 3507  
QY 4982 TACCGGCACTGTACCAACCTTACCTTACCTGCGGCTTACCCCGACAGCGGCGGCTGGAG 5041  
DB 3508 TACCGGCACTGTACCAACCTTACCTTACCTGCGGCTTACCCCGACAGCGGCGGCTGGAG 3567

QY 5042 AACCGGACAGCATCATGATGATCATCATCTTCGACGACGATGACCAACACAGCGCC 5101  
DB 3568 AACCGGACAGCATCATGATGATCATCATCTTCGACGACGATGACCAACAGCGGCC 3627  
QY 5102 ACCGCAATGGGCCGAGCGAGCTGATATGCTGAGGGGGCTCTGGCCCCGAGGCTCGGCTG 5161  
DB 3628 ACCGCAATGGGCCGAGCGAGCTGATATGCTGAGGGGGCTCTGGCCCCGAGGCTCGGCTG 3687  
QY 5162 GCACTCAACTAGCTGCGGGGTCCCGAGGACATCATGACCTGTGCCCAAGTGCACACCTG 5221  
DB 3688 GCACTCAACTAGCTGCGGGGTCCCGAGGACATCATGACCTGTGCCCAAGTGCACACCTG 3747  
QY 5222 CCTGTGCTGTGCCCCGACACACAGGACCCCGACCTCGCATGAGACCGCTTGGCTTAC 5281  
DB 3748 CCTGTGCTGTGCCCCGACACACAGGACCCCGACCTCGCATGAGACCGCTTGGCTTAC 3807  
QY 5282 CTCCCCACCGGCCCCCAACCCCTTACAGACCGGACACAGAGCTCCCACTGCCCAAGA 5341  
DB 3808 CTCCCCACCGGCCCCCAACCCCTTACAGACCGGACACAGAGCTCCCACTGCCCAAGA 3867  
QY 5342 GATCCACACACTTGAACAAAACCAACCAACAGCTCTGTCCGAGCGGAGGAGACCGG 5401  
DB 3868 GATCCACACACTTGAACAAAACCAACCAACAGCTCTGTCCGAGCGGAGGAGACCGG 3927  
QY 5402 GATGAGAGCGGAGACCGGAGATCGGAGCGGGAAGTCCATCTCTTCACTCCACACGAC 5461  
DB 3928 GATGAGAGCGGAGACCGGAGATCGGAGCGGGAAGTCCATCTCTTCACTCCACACGAC 3987  
QY 5462 GTGAGGACGACCCCATGTGAGGACCTGTGACAGAGCAGAGAGCGGACGCGGACG 5521  
DB 3988 GTGAGGACGACCCCATGTGAGGACCTGTGAGGACATGAGAGCTCCACCCCGGCTCG 4047  
QY 5522 AGCGCGGAGGTGAGGAGCAGACGACGCGCCCGCTCCACCTCCATGCCACAGAC 5581  
DB 4048 GGACTCTTGTGAGGCGGAGAGAGGCTCCCGTGTGATGACATGACATGACACAGAC 4107  
QY 5582 TGGCCATCTCCCTCGGACCCAGAGATGCTTCCAGCAGAGACCTTGTATACAC 5641  
DB 4108 TGGCCATCTCCCTCGGACCCAGAGATGCTTCCAGCAGAGACCTTGTATACAC 4167  
QY 5642 ACAAGCAGAAAGGATATACACGCTGTGAGAGCCCAAGGCCACAGGATCCGATGAGTCC 5701  
DB 4168 ACAAGCAGAAAGGATATACACGCTGTGAGAGCCCAAGGCCACAGGATCCGATGAGTCC 4227  
QY 5702 ACTTCACCTCTCAACCGGTTGCGCCAGCTGCAATTTCCACCTTGCCACCTGCCA 5761  
DB 4228 ACTTCACCTCTCAACCGGTTGCGCCAGCTGCAATTTCCACCTTGCCACCTGCCA 4287  
QY 5762 CTGGCGGACCCCTTGATGGGGTCTTACCTTACCTCATGAGACCGGTCTTGTGCGCAAG 5821  
DB 4288 CTGGCGGACCCCTTGATGGGGTCTTACCTTACCTCATGAGACCGGTCTTGTGCGCAAG 4347  
QY 5822 GAGGCCCCCGGGTTCGCGCGGCAAGAGGGCCCGAGAGACACCGGCAATGCTTTCG 5881  
DB 4348 GAGGCCCCCGGGTTCGCGCGGCAAGAGGGCCCGAGAGACACCGGCAATGCTTTCG 4407  
QY 5882 GCCAAGCCCCAGCGCGCTGCGGCTGAGAGCCGCTCTCCCGCAGCAAGGGCTCGAG 5941  
DB 4408 GCCAAGCCCCAGCGCGCTGCGGCTGAGAGCCGCTCTCCCGCAGCAAGGGCTCGAG 4467  
QY 5942 CCGGCGCCCTTATGATGCTCTGTCTTGGCCACGCAACATGCGCCGACCCCTGCGAAG 4527  
DB 4468 CCGGCGCCCTTATGATGCTCTGTCTTGGCCACGCAACATGCGCCGACCCCTGCGAAG 4587  
QY 6002 AACCTGACACCTTACAGGACCGCGGACCGCGCGGACCACTGAGGCTCGGAGC 6061  
DB 4528 AACCTGACACCTTACAGGACCGCGGACCGCGCGGACCACTGAGGCTCGGAGC 4587  
QY 6062 CCGACCGGAAAGACTCAAGTAAACCTTTTCATTCAGAGAACTGAACTCGGTTCT 6121  
DB 4588 CCGACCGGAAAGACTCAAGTAAACCTTTTCATTCAGAGAACTGAACTCGGTTCT 4647  
QY 6122 CTG----- 6124

Db 4648 CTGGGTAAAGCAACCTGACAGGCGCACTTATAGACGGATATCATGCGTCAATT 4707  
6125 -----GCT 6127  
Db 4708 GCTCAGATTAAGGGGGCGGAGAGAGGTGGCCGTGGCCACGAGCTCCCTCGCGATGGT 4767  
6128 TACACGGGAGCAGCTTACAGCCCCGAAAGGGGTGAGCCCGTCAAGCCCTGTGAGCTCAACC 6187  
4768 TACACGGGAGCAGCTTACAGCCCCGAAAGGGGTGAGCCCGTCAAGCCCTGTGAGCTCAACC 4827  
6188 AGTCTGACCCACGACAAAGGGGCTCCCAAGACCTGGAAGAGCTCGAACAAGCCACTTG 6247  
4828 AGTCTGACCCACGACAAAGGGGCTCCCAAGACCTGGAAGAGCTCGAACAAGCCACTTG 4887  
6248 GAGGGGAGAGTGGGGCCCAAGGAGCCGAGCCCGTGAAGCTTGGCGGGAGGCGCCAC 6307  
4888 GAGGGGAGAGTGGGGCCCAAGGAGCCGAGCCCGTGAAGCTTGGCGGGAGGCGCCAC 4947  
6308 CTCCACACCTTGGGCGCGCTGCTGAGAGAGCCAGCCCTGCTCAAGCCCGCTCCAGACC 6367  
4948 CTCCACACCTTGGGCGCGCTGCTGAGAGAGCCAGCCCTGCTCAAGCCCGCTCCAGACC 5007  
6368 GCCCAGGGGCTCAAGGTCAACAGCGGTGTCAACCTGAGCCAGCAATCACT----- 6421  
5008 GCCCAGGGGCTCAAGGTCAACAGCGGTGTCAACCTGAGCCAGCAATCACTGTAAC 5067  
6422 ----- 6421  
5068 AGCGCTCTCTGCTGCTGTGCTTCACTTGTGACCTGAGGGGACAGAGCCTGAGAGGGG 5127  
6422 ----- 6421  
5128 ATGGGGAACCCACAGCCCTTCTGTCTGCGGGGGTGGCTGGGGGATCCAGGGCATGGGG 5187  
6422 -----GAGGTATCAACAGAGACTTACCCGAGCCAGCCAG 6460  
5188 CTGGGGGAGATCCAGGGCGTGGGAGGTCAATCAACAGAGCTACCCGGGACCAACCCAG 5247  
6461 CAGCTCAAGGAGCAGCCCTGCGCGCGCTTACTCTTCTGCGGGGCAAGTCCCGCTC 6520  
5248 CAGCTCAAGGAGCAGCCCTGCGCGCGCTTACTCTTCTGCGGGGCAAGTCCCGCTC 5307  
6521 CTGACCTCCGCGGCCACCGAGTGACTTACTCCCGCGCGGACCATGGTGGCCCG 6580  
5308 CTGACCTCCGCGGCCACCGAGTGACTTACTCCCGCGCGGACCATGGTGGCCCG 5367  
6581 GCCCGTGGCTCCCGCCACAGCGAAGGGGGCAAGGTCTTCAAGGCCAACAAGACGTGG 6640  
5368 GCCCGTGGCTCCCGCCACAGCGAAGGGGGCAAGGTCTTCAAGGCCAACAAGACGTGG 5427  
6641 GTCTTGGGTGGTGGTGAAGACGGTATTGAACCTGTGTCCCAACGAGGAGGATGACGAG 6700  
5428 GTCTTGGGTGGTGGTGAAGACGGTATTGAACCTGTGTCCCAACGAGGAGGATGACGAG 5487  
6701 CAGGGGCACTCCCGAGTGTGTGATCCGCGCTGTACCGGGATGGGGAAAGAGAGGAG 6760  
5488 CAGGGGCACTCCCGAGTGTGTGATCCGCGCTGTGTACCGGGATGGGGAAAGAGAGGAG 5547  
6761 CCAGAGGATGGGCTCAAGTCTCCAGGCAACAGGACGAGCGGCAAGCTTCTTCAAG 6820  
5548 CC---CAGGATGGGCTCAAGTCTCCAGGCAACAGGACGAGCGGCAAGCTTCTTCAAG 5604  
6821 AAGCTGAACGGAAGCACTCCCGCATGTGTCAAGTGTCAAGAGCAAGATCAACAGAG 6880  
5605 AAGCTGAACGGAAGCACTCCCGCATGTGTCAAGTGTCAAGAGCAAGATCAACAGAG 5664  
6881 CTGAACACCAACAACCGGATGAGCTGAATCAATATCAAGCGAGCTGGAGCGAGATC 6940  
5665 CTGAACACCAACAACCGGATGAGCTGAATATCAATATCAAGCGAGCTGGAGCGAGATC 5724  
6941 TTCAATATGCCCCCATCAACCGGACAGGCTTATGACCTATGAAGCAAGCGGCTGCA 7000

Db 5725 TTCAATATGCCCCCATCAACCGGACAGGCTTATGACCTATGAAGCAAGCGGCTGCA 5784  
7001 GAACATGCGAGCAACAATGGGGCTGAGGGCCATTAATGAAGGACCTATGGGTAA 7060  
5785 GAACATGCGAGCAACAATGGGGCTGAGGGCCATTAATGAAGGACCTATGGGTAA 5844  
7061 TATGACCAAGTGGAGAGTCCCGCGCTCAAGCGCAATCTTTTAACTCTGTAATGCC 7120  
5845 TATGACCAAGTGGAGAGTCCCGCGCTCAAGCGCAATCTTTTAACTCTGTAATGCC 5904  
7121 AGTCCAGCTCCCGCTGCTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 7180  
5905 AGTCCAGCTCCCGCTGCTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 5964  
7181 CTCACTGCGAGAGTGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7240  
5965 CTCACTGCGAGAGTGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6024  
7241 GCCAAGTCCCGCGCGCGCGCTGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7300  
6025 GCCAAGTCCCGCGCGCGCGCGCTGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6084  
7301 CACTCGAGAGGAGAGTCAACCGCGGAGCGGCTCAACCAACCGGCTGAGAGAGAGG 7360  
6085 CACTCGAGAGGAGAGTCAACCGCGGAGCGGCTCAACCAACCGGCTGAGAGAGAGG 6144  
7361 CCTCGTCCGAGAGTTCACAGCCATTCCTTACAAACCTCTGATCAATGCGGCTGAGAGG 7420  
6145 CCTCGTCCGAGAGTTCACAGCCATTCCTTACAAACCTCTGATCAATGCGGCTGAGAGG 6204  
7421 GGTGTATGAGCTTCCCAACCGGAGCTCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 7480  
6205 GGTGTATGAGCTTCCCAACCGGAGCTCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 6264  
7481 CCCACCAAGCGCTGGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7540  
6265 CCCACCAAGCGCTGGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6324  
7541 TCCGACAGGAGTGA 7555  
6325 TCCGACAGGAGTGA 6339

RESULT 11  
BC004326 2842 bp mRNA linear PRI 12-JUL-2001  
LOCUS BC004326  
DEFINITION Homo sapiens, clone IMAGE:3629785, mRNA, partial cds.  
ACCESSION BC004326  
VERSION BC004326.1 GI:13279241  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAR-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK  
COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@gscc.bc.ca](mailto:info@gscc.bc.ca)  
Steven Jones, Jennifer Asano, Ian Bowdler, Yaron Butterfield,  
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,

Leticia Heiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeary, Steven Nees, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Serier, Diane Smalins, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Teai, Nataaja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Zetter.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILM, at: <http://image.llnl.gov> Series: IRL. Plate: 13 Row: i Column: 14 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES  
source

Location/Qualifiers  
1..2842  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3629785"  
/tissue\_type="Placenta", choriocarcinoma"  
/clone\_lib="NIH MGC\_21"  
/lab\_host="DH10B-R"  
/note="vector: pOTB7"  
<1..1821  
/codon\_start=1  
/product="Unknown (protein for IMAGE:3629785)"  
/protein\_id="AAH04326.1"  
/db\_xref="GI:13279242"  
/translation="OAGVMSPLGSDTLDVYPTIMEVLPKKAPRYAPRRPADT  
GHAFAPKAPASGLEBPASPSKGSFRLVLPVPSGHATATPAKMLPHHASPAPPA  
PPASASDHRERKTSKPSIQLELRSLGTHSSYSPPGVSVSSPSLTHDKGLP  
KHLELDKSHLEGLRKPQPGVKIGRAHMLPLRPBESQSSPLQTPAGVKG  
ORVTLQHISEVITODYTRHNPOLASPLPAPLYSPGASCPVLDLRPSDYLPLP  
PDHGAAPASGHSSEKGRSPENKTSVVGSGDGLRPSPEGMTEPGHSSAVYPL  
YRDRORPSPMSGNTSOPRAPEKLTESNMAKSKKOKNTKNTLNHNEPE  
YNISOPGTEIRNPAITGTMTYRSQAVQEHASTNMLBLIRKALMGKTDQNEBP  
PLSAAFPPLNASSALPLAMPPLTAADGSDHTLSPGGGAKAKVSGRSSKASAP  
GLASGDRPVSYSVSBGDCNRKRTPLTRVWEDRSPFAGSTPPFPNPLIMLOAGVMA  
SPPPGLPAGSGPLAGPHHAWDEBPKPLCQYETLSDSE"

ORIGIN

Query Match 32.2%; Score 2759.8; DB 8; Length 2842;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 2798; Conservative 2; Mismatches 15; Indels 12; Gaps 2;

5746	TGCCACCACTGCGGCGGACACCTGATGGGGGTCTACCTCATCGATGAGCC	5805
12	TGTCTGCGCTCCCACTGGGCGGACCTCGATGGGGGTCTACCTCATGAGCC	71
5806	CGTCTTGTGCGCAAGAGAGCCCGGGGTGCGCCGCGCAAGACGCGCCGAGCAGAC	5865
72	CGTCTTGTGCGCAAGAGAGCCCGGGGTGCGCCGCGCAAGACGCGCCGAGCAGAC	131
5866	CGGCGATGCTTCTCTGCGCAAGCCCGCGCGCTCGGGCTGAGACCGCGCTCTCC	5925
132	CGGCGATGCTTCTCTGCGCAAGCCCGCGCGCTCGGGCTGAGACCGCGCTCTCC	191
5926	CAGCAAGGCTCTGAGCCCGCGCGCTGATGCTCTGCTCTGCGCAAGCAGCATGCG	5985
192	CAGCAAGGCTCTGAGCCCGCGCGCTGATGCTCTGCTCTGCGCAAGCAGCATGCG	251
5986	CGGCAAGGCTCTGAGCCCGCGCGCTGATGCTCTGCTCTGCGCAAGCAGCATGCG	6045
252	CGGCAAGGCTCTGAGCCCGCGCGCTGATGCTCTGCTCTGCGCAAGCAGCATGCG	311
6046	TGCTCTGCGCTCTGAGCCCGCGCAAGCAAGTAAACCTTTTTCATCGAGCA	6105
312	TGCTCTGCGCTCTGAGCCCGCGCAAGCAAGTAAACCTTTTTCATCGAGCA	371
6106	ACTGGAACCTGCTTCTGAGGTTACCAAGGCGAGCATACAGCCCGGAAAGGGGTGAGCC	6165
372	ACTGGAACCTGCTTCTGAGGTTACCAAGGCGAGCATACAGCCCGGAAAGGGGTGAGCC	431

QY	6166	CGTCAGCCCTGTGAGCTTACCCAGTGTGACCCAGCAGCAGGAGGCTCTCCCAAGCATTGGA	6225
DB	432	CGTCAGCCCTGTGAGCTTACCCAGTGTGACCCAGCAGCAGGAGGCTCTCCCAAGCATTGGA	491
QY	6226	AGAGCTGCAAGAGCCACTGAGAGGGAGCTGGGCGCCAGAGCAGGAGCCCGGTGA	6285
DB	492	AGAGCTGCAAGAGCCACTGAGAGGGAGCTGGGCGCCAGAGCAGGAGCCCGGTGA	551
QY	6286	GCTTGGCGGGAGGCGCCCACTTCCCAACCTGGCGCGCTGCTTGAAGCAGCCCTC	6345
DB	552	GCTTGGCGGGAGGCGCCCACTTCCCAACCTGGCGCGCTGCTTGAAGCAGCCCTC	611
QY	6346	GTCACGCCCTGCTTCCAGACCGGCCAGGGGTCAAGGTCAACGCGGTGTCACTCT	6405
DB	612	GTCACGCCCTGCTTCCAGACCGGCCAGGGGTCAAGGTCAACGCGGTGTCACTCT	671
QY	6406	GGCCAGAGCAATCATGAGGTCTATCAACAGAGTCAACCGGGACCCAGCAGCAGCT	6465
DB	672	GGCCAGAGCAATCATGAGGTCTATCAACAGAGTCAACCGGGACCCAGCAGCAGCT	731
QY	6466	CAGGCAACCCCTGCGCGCCCTCTACTCTTCCCTTGGGCGCAGCTGCCGCTCTGA	6525
DB	732	CAGGCAACCCCTGCGCGCCCTCTACTCTTCCCTTGGGCGCAGCTGCCGCTCTGA	791
QY	6526	CCTCGCGCCCAACCCAGTACTTCTACCTCCCGCCCGGACCAAGTGTCCCGGCGG	6585
DB	792	CCTCGCGCCCAACCCAGTACTTCTACCTCCCGCCCGGACCAAGTGTCCCGGCGG	851
QY	6586	TGCTCTCCCGCCCAACCGGAGGGGGCAAGAGGTCTCCAGGCAACAGAGGTGTCTT	6645
DB	852	TGCTCTCCCGCCCAACCGGAGGGGGCAAGAGGTCTCCAGGCAACAGAGGTGTCTT	911
QY	6646	GGGTGTGTGAGGACGATATTGAACCTGTGTCCCAACCGGAGGCGATGACGAGCCAG	6705
DB	912	GGGTGTGTGAGGACGATATTGAACCTGTGTCCCAACCGGAGGCGATGACGAGCCAG	971
QY	6706	GOACTCCGAGGTGTGTACCGGCTGTCTGATCCGGATGGGAAACAGAGAGCCAG	6765
DB	972	GOACTCCGAGGTGTGTACCGGCTGTCTGATCCGGATGGGAAACAGAGAGCCAG	1031
QY	6766	CAGATGGGCTTCAAGTCTCAGGCAACAGCAGCGCGCAAGCTTCTCAGCAAGCT	6825
DB	1032	CAGATGGGCTTCAAGTCTCAGGCAACAGCAGCGCGCAAGCTTCTCAGCAAGCT	1091
QY	6826	GACCGAGCAACTCCGCGATGCTCAAGTCCCAAGAGCAAGATCAACAGAACTGA	6885
DB	1092	GACCGAGCAACTCCGCGATGCTCAAGTCCCAAGAGCAAGATCAACAGAACTGA	1151
QY	6886	CACCCACAAACCGAATGAGCTTAATATCAATATCAAGCAGCCTGGGACGAGATTTCA	6945
DB	1152	CACCCACAAACCGAATGAGCTTAATATCAATATCAAGCAGCCTGGGACGAGATTTCA	1211
QY	6946	TATGCGGCGCATCAACCGGAAGGCTTAATGAGCAAGTATGAGGCGGAGGAGAA	7005
DB	1212	TATGCGGCGCATCAACCGGAAGGCTTAATGAGCAAGTATGAGGCGGAGGAGAA	1271
QY	7006	TGCGAGCAACAATGGGGCTGAGAGCCATTAATTAAGAAAGCACTCATGGTAAATATGA	7065
DB	1272	TGCGAGCAACAATGGGGCTGAGAGCCATTAATTAAGAAAGCACTCATGGTAAATATGA	1331
QY	7066	COAGTGGGAAGATCCCGCGCTCAGCGCAATGCTTTTAACTCTGAATGCCAGTGC	7125
DB	1332	COAGTGGGAAGATCCCGCGCTCAGCGCAATGCTTTTAACTCTGAATGCCAGTGC	1391
QY	7126	CAGCTGCGCGCTGATAGCCATTAACCGGCTGAGAGGAGAGGTGACCAACCTGAC	7185
DB	1392	CAGCTGCGCGCTGATAGCCATTAACCGGCTGAGAGGAGAGGTGACCAACCTGAC	1451
QY	7186	CTGCGCAGGTGGCGGGAGGCAAGGTCTTGGCGAGACCCAGCAGCCGAAAGCCAA	7245
DB	1452	CTGCGCAGGTGGCGGGAGGCAAGGTCTTGGCGAGACCCAGCAGCCGAAAGCCAA	1511
QY	7246	GTCGCCGCGCCCGGCGCTGAGCATCTGGGAGCCGCGCACCTCTGTCTCTCATGTCATCT	7305

Db 1512 GTCCCGGCGCCCGGCGCTGGCATCTGGGGAGCCGGCCACCTCTGTCTCTCACTGACATCG 1571  
Qy 7306 GGAGGGAGACTGCAACCGCGCGAGCGCGCTCAACCAACCGGCTGGAGAGAGAGCGCCCTC 7365  
Db 1572 GGGGGAGAGCTGCAACCGCGCGAGCGCGCTCAACCAACCGGCTGGAGAGAGAGCGCCCTC 1631  
Qy 7366 GTCCGAGAGTTCCACCGCATCTCCCTCAACACCCCTGATCATGGCTGACAGCGGGTGT 7425  
Db 1632 GTTCGAGAGTTCCACCGCATCTCCCTCAACACCCCTGATCATGGCTGACAGCGGGTGT 1691  
Qy 7426 CATGCTTCCCGACCCCGGCGCTCCCGCGGAGCGGGGCGCCCTGCTGCGCCCGCA 7485  
Db 1692 CATGCTTCCCGACCCCGGCGCTCCCGCGGAGCGGGGCGCCCTGCTGCGCCCGCA 1751  
Qy 7486 CCAAGCGCTGGAGAGAGAGAGCGCCACCTGCTCTGCTGCAATAGAGACATCTCCCA 7545  
Db 1752 CCAAGCGCTGGAGAGAGAGAGCGCCACCTGCTCTGCTGCAATAGAGACATCTCCCA 1811  
Qy 7546 CAGCGAGTACTCAGAACAGAGCGGGGGGGGGGGCGGCGGTGEBRSPRLUGTCAGGTC 7605  
Db 1812 CAGCGAGTACTCAGAACAGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1860  
Qy 7606 ACCGAGCCAAGAGAGCGGCGCTGCAAGAGCGGGGGGGGGGGGGGGGGGGGGGGGG 7665  
Db 1861 ACCGAGCCAAGAGAGCGGCGCTGCAAGAGCGGGGGGGGGGGGGGGGGGGGGGGGG 1920  
Qy 7666 AGGAGCGGCGTGGAGTCCGCGCTGCGCGCTCCATCTGCTCTGCTCAAGCGCGCATCT 7725  
Db 1921 AGGAGCGGCGTGGAGTCCGCGCTGCGCGCTCCATCTGCTCTGCTCAAGCGCGCATCT 1980  
Qy 7726 CCTGTCTAAGCGCTTAAGTCTCCGCGCGCGGGCGCGCTGCTGAGACCTTACTCA 7785  
Db 1981 CCTGTCTAAGCGCTTAAGTCTCCGCGCGCGGGCGCGCTGCTGAGACCTTACTCA 2040  
Qy 7786 GGGGAGTGTTCCTGTGCTCGGGAAGGAGGGGAGGGGCGGAGGGGCGGAGCGGCA 7845  
Db 2041 GGGGAGTGTTCCTGTGCTCGGGAAGGAGGGGAGGGGCGGAGGGGCGGAGCGGCA 2100  
Qy 7846 GCGTGTGAGAGCCACACAGAGCGGCGGAGGGGCGGAGGGGCGGAGGGGCGGAGCG 7905  
Db 2101 GCGTGTGAGAGCCACACAGAGCGGCGGAGGGGCGGAGGGGCGGAGGGGCGGAGCG 2160  
Qy 7906 GCACCTCCAGCGCATGCGCTCCCGCGGAGTGTGGAACCAAGCTTAAGTGAAGCTG 7965  
Db 2161 GCACCTCCAGCGCATGCGCTCCCGCGGAGTGTGGAACCAAGCTTAAGTGAAGCTG 2220  
Qy 7966 CAGCCCGCGCGCGCTCCCGCGCTCCAGTCCGCTTGAAGCGCTTGAAGAGATGAGCG 8025  
Db 2221 CAGCCCGCGCGCGCTCCCGCGCTCCAGTCCGCTTGAAGCGCTTGAAGAGATGAGCG 2280  
Qy 8026 GCGCGCTGTCCAGCGCGCGCGCGCGCTGCTTCCGGTCCCGCAAGCTGCCCGCAAGCG 8085  
Db 2281 GCGCGCTGTCCAGCGCGCGCGCGCGCTGCTTCCGGTCCCGCAAGCTGCCCGCAAGCG 2340  
Qy 8086 ATTTGCTGGAACCAAGTCAAGCGAGGTGGAGCGGAGCAAAAGGGGCGAGTGGCGCTG 8145  
Db 2341 ATTTGCTGGAACCAAGTCAAGCGAGGTGGAGCGGAGCAAAAGGGGCGAGTGGCGCTG 2400  
Qy 8146 GAAAGGATGCTCCGAGAGCTGGAATGTTTTTTCACACATCGTTGCGGAGCGGTGGAA 8205  
Db 2401 GAAAGGATGCTCCGAGAGCTGGAATGTTTTTTCACACATCGTTGCGGAGCGGTGGAA 2460  
Qy 8206 GGAAGGAGGAGTGAATGATGTTGTTTCAAGGGTATATTTTGAATCTTCAATGA 8265  
Db 2461 GGAAGGAGGAGTGAATGATGTTGTTTCAAGGGTATATTTTGAATCTTCAATGA 2520  
Qy 8266 ATTTAATTCAGATGTTTTCAGCAAGAGAGCTTACCAGATTTAAGTGTGCTGCTTT 8325  
Db 2521 ATTTAATTCAGATGTTTTCAGCAAGAGAGCTTACCAGATTTAAGTGTGCTGCTTT 2580  
Qy 8326 GATCTCTGCTTACCGTTCAAGAGCGGTGTCAGGCGGAGAGTGGATCCCATCTCG 8385

Db 2581 GATCTCTGCTTACCGTTCAAGAGCGGTGTCAGGCGGAGAGTGGATCCCATCTCG 2640  
Qy 8386 CAGAACCAAGGGGGGGGGAGCTGCTCG-TCAGCGCGCGCTGTGTCCTCCCTCCCTC 8444  
Db 2641 CAGAACCAAGGGGGGGGGAGCTGCTCGCTCAGCGCGCGCTGTGTCCTCCCTCCCTC 2700  
Qy 8445 CTTTGGGAGAGTGAATTCGATGCGTATCTGTGCGCGCCATTTTGGCAGAGGTGTGTA 8504  
Db 2701 CTTTGGGAGAGTGAATTCGATGCGTATCTGTGCGCACCATTTTGGCAGAGGTGTGTA 2760  
Qy 8505 TTCTGTATTTTACACAGCTGCTCTTAATTAAGAGATTAATTAATTAATTAATTA 8564  
Db 2761 TTCTGTATTTTACACAGCTGCTCTTAATTAAGAGATTAATTAATTAATTAATTA 2820  
Qy 8565 AAAAAA 8571  
Db 2821 AAAAAA 2827

RESULT 12  
S83390  
LOCUS 2930 bp mRNA linear PRI 28-MAR-1997  
DEFINITION T3 receptor-associating cofactor-1 [human, fetal liver, mRNA, 2930 nt].  
ACCESSION S83390  
VERSION S83390.1 GI:1911769  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 2930)  
Sande, S. and Privalsky, M.L.  
Identification of TRACs (T3 receptor-associating cofactors), a family of cofactors that associate with, and modulate the activity of, nuclear hormone receptors  
Mol. Endocrinol. 10 (7), 813-825 (1996)  
JOURNAL Genbank staff at the National Library of Medicine created this PUMED entry [NCBI glibbeq 179975] from the original journal article.  
REMARK Location/Qualifiers  
FEATURES  
source  
1..2930  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..2930  
/gene="T3 receptor-associating cofactor-1, TRAC-1"  
456..2765  
/gene="T3 receptor-associating cofactor-1, TRAC-1"  
/note="Authors report detecting several TRAC alternative transcripts; TRAC-1"  
/codon\_start=1  
/product="T3 receptor-associating cofactor-1"  
/protein\_id="AAB50847.1"  
/db\_xref="GI:1911770"  
/rnelation="MAORADMRGLSPRESSALNYAAGPGIIDSOUPLVYVP  
TGTPTATMDRLAYLPAPQPPSRHSSPSLPQGTTHTKTTTSSSRBRDRBR  
DDREREXSILSTTVVHAPVWRPTEGSSSSGSSGGGSSSRPAASHAHQHP  
ISPRQDALQORPSVLTHTGMKGIITAVPSPPTVLRSTSSPVAPATFPATCP  
LGSTLDGVPTLMEVPLVKEAPRVARPERPADTGATGAPASGEPASPSKG  
SEPRPLVPVSGHATITARTPAKNLA.PHNASPPRPAAPASDPHREKTSKPSIGEL  
ELRSIGYSSSPGVSPVSPVSPSLTHDGLRPHLELSHLEGLRPPQOPRV  
KLGEERAHLPRLPSPSSPILQTRPVYGHORVYTLAQHISEVYTDYTRHP  
QOLSAVPLVSPFPGASCVIDLRPBDLVLPDRHAPASGSHSGKSPBN  
KTSVAGGEGDGLVPSPEGMTEPGHRSRAVPLVLRDQOYRPSHSGKSPBN  
PAFSPKLTESNAMYKSKOJINKLTNRMEPEYNISQPTBLIFNMAI1GCTGIMT  
YRSQAVOEHASTNMGLBALIRKALMGGAGKAVSGRPSRKAQSPAPGASGRPSV  
SSVHSGCNRPTPLTNVWEDRPSASGTPPNVPLIMRLAQAWASPPPGGLRPS  
GPLAGNHAHMBRPPPLCSQYETYSDBS"

ORIGIN  
Query Match 30.6%; Score 2618.6; DB 8; Length 2930;



Best Local Similarity 94.6%; Pred. No. 0; Matches 2786; Conservative 3; Mismatches 9; Indels 146; Gaps 2;									
QY	4785	CCCCGACAGCAGCCGTCGCCGAGCAACACCCACATCTCGCCCTATGAGCACTGCG	4844						
Db	133	CCCCGACAGCAGCCGTCGCCGAGCAACACCCACATCTCGCCCTATGAGCACTGCG	192						
QY	4845	TTCCGGGGGTGAGTGGCTGGAACCTGTATCGAGCCACATCCGCCCTTGACCCCA	4904						
Db	193	TTCCGGGGGTGAGTGGCTGGAACCTGTATCGAGCCACATCCGCCCTTGACCCCA	252						
QY	4905	CCTCCATACCCCGGCGCATCTCTTGAGAGCAGCCGCTGCTTAACCTTGCCCGACAC	4964						
Db	253	CCTCCATACCCCGGCGCATCTCTTGAGAGCAGCCGCTGCTTAACCTTGCCCGACAC	312						
QY	4965	TGGCCCCCAACCCCACTTACCCGCACTGTACCCACCCCTACCTCATCGCGGCTACCCCG	5024						
Db	313	TGGCCCCCAACCCCACTTACCCGCACTGTACCCACCCCTACCTCATCGCGGCTACCCCG	372						
QY	5025	ACACGGCGCGCTGAGAAACCGGACAGCCATCATGATGATCACTTGCAGACAGAGA	5084						
Db	373	ACACGGCGCGCTGAGAAACCGGACAGCCATCATGATGATCACTTGCAGACAGAGA	432						
QY	5085	TGCACCAACAACGGCCACCCGCACTGAGCCCAAGAGCTGATTTGCTGAGGGGCTCTCGC	5144						
Db	433	TGCACCAACAACGGCCACCCGCACTGAGCCCAAGAGCTGATTTGCTGAGGGGCTCTCGC	492						
QY	5145	CCCGGAGCTCTCGTGGCACTCAACTAGCTGCGGGTCCCGGAGCATCATGACCTGT	5204						
Db	493	CCCGGAGCTCTCGTGGCACTCAACTAGCTGCGGGTCCCGGAGCATCATGACCTGT	552						
QY	5205	CCCAAGTGCCACACTGCTGTGTCTGTGCTCCCGACACAGGCAACCCAGGCCACCGCA	5264						
Db	553	CCCAAGTGCCACACTGCTGTGTCTGTGCTCCCGACACAGGCAACCCAGGCCACCGCA	612						
QY	5265	TGGACCGCTTTCCTTACCTTCCCAACGCGCCCAAGCTTTCAGCAGCCGCGCACAGAGCT	5324						
Db	613	TGGACCGCTTTCCTTACCTTCCCAACGCGCCCAAGCTTTCAGCAGCCGCGCACAGAGCT	672						
QY	5325	CCCACTCTCCCGAGAGGATCCAAACAATTGACAAAACAACAACAACAGTCTGTGCG	5384						
Db	673	CCCACTCTCCCGAGAGGATCCAAACAATTGACAAAACAACAACAACAGTCTGTGCG	732						
QY	5385	AGCGGAGCGAGACCGGATTCAGAGCGGGAACCGGATCGGAGCGGGAAGTCCATCC	5444						
Db	723	AGCGGAGCGAGACCGGATTCAGAGCGGGAACCGGATCGGAGCGGGAAGTCCATCC	792						
QY	5445	TCAAGTCCACACACGATGAGAGCAACGCACTTGAAGACCTGTGTAACAAGCAGAGCA	5504						
Db	793	TCAAGTCCACACACGATGAGAGCAACGCACTTGAAGACCTGTGTAACAAGCAGAGCA	852						
QY	5505	GGGGAGAGCGGAGCGGAGCGGCGGGGGTGGGGGAGCAGACGCGCCCGCTCCACT	5564						
Db	853	GGGGAGAGCGGAGCGGAGCGGCGGGGGTGGGGGAGCAGACGCGCCCGCTCCACT	912						
QY	5565	CCCATGCCCCACGAGACTCGCCCATCTCCCTCGAGACCGAGATGCTTCAGACAGAGAC	5624						
Db	913	CCCATGCCCCACGAGACTCGCCCATCTCCCTCGAGACCGAGATGCTTCAGACAGAGAC	972						
QY	5625	CCAGTGTGCTTCAACAACAAGGATGAGAGGATATCAACCGCTGTGAGCCCAAGCAGC	5684						
Db	973	CCAGTGTGCTTCAACAACAAGGATGAGAGGATATCAACCGCTGTGAGCCCAAGCAGC	1032						
QY	5685	CCAGGATGCTGAGGTCACCTCAACCTCTCAACCGGTTGGCCAGGTGCAATTCCAC	5744						
Db	1033	CCAGGATGCTGAGGTCACCTCAACCTCTCAACCGGTTGGCCAGGTGCAATTCCAC	1092						
QY	5745	CTGCAACCACTGACCACTGGGCGGCAACCTCGATGGGGTCTAACCTTCACTGAGAC	5804						
Db	1093	CTGCAACCACTGACCACTGGGCGGCAACCTCGATGGGGTCTAACCTTCACTGAGAC	1152						
QY	5805	CGGTCTTGCTGCGCAAGAGGCCCCCGGGTGGCCCGGCAGAGCGGCCCAAGACACA	5864						

Db	1153	CGGTCTGCTGCGCAAGAGGCCCCCGGGTGGCCCGGCAGAGCGGCCCAAGACACA	1212						
QY	5865	CCGACATGCTCTCTCGCCAAAGCCGCCAGCCGCTCGGGCTGAGGCCGCTCTCC	5924						
Db	1213	CCGACATGCTCTCTCGCCAAAGCCGCCAGCCGCTCGGGCTGAGGCCGCTCTCC	1272						
QY	5925	CCAGCAAGGGCTCGAGGCCCGGCTTATGTCCTTCTGTCTTGAGCCACGCAACCATG	5984						
Db	1273	CCAGCAAGGGCTCGAGGCCCGGCTTATGTCCTTCTGTCTTGAGCCACGCAACCATG	1332						
QY	5985	CCCGACCCCTGCGAAGAACCTCGACCTCAACAAGCCAGGCCCGGACCGCGGGCCAC	6044						
Db	1333	CCCGACCCCTGCGAAGAACCTCGACCTCAACAAGCCAGGCCCGGACCGCGGGCCAC	1392						
QY	6045	CTGCTCGGCTCGAGACCGGCAACCGGAAAGACTCAAGTAAACCTTTTCATCGAG	6104						
Db	1393	CTGCTCGGCTCGAGACCGGCAACCGGAAAGACTCAAGTAAACCTTTTCATCGAG	1452						
QY	6105	AACTGAACTCGGTTCTTGAGTTACACGGGACAGCTTACAGCCCGGAGGGTGAAC	6164						
Db	1453	AACTGAACTCGGTTCTTGAGTTACACGGGACAGCTTACAGCCCGGAGGGTGAAC	1512						
QY	6165	CCGTAGCCCTGTGAGCTTCAACCACTGACCCAGACAGAGGGGCTCCCAAGACCTGG	6224						
Db	1513	CCGTAGCCCTGTGAGCTTCAACCACTGACCCAGACAGAGGGGCTCCCAAGACCTGG	1572						
QY	6225	AAGACTGCAACAAGACCACTTGAGGGGAGCTGCGGCCAAGACCGACCGCCGTGA	6284						
Db	1573	AAGACTGCAACAAGACCACTTGAGGGGAGCTGCGGCCAAGACCGACCGCCGTGA	1632						
QY	6285	AGCTTGGGGGGAGGCGGCCACTTCCCAACCTGCGCGCTGCTGAGAGCCAGCCTT	6344						
Db	1633	AGCTTGGGGGGAGGCGGCCACTTCCCAACCTGCGCGCTGCTGAGAGCCAGCCTT	1692						
QY	6345	CGTCAAGCCGCTGTCCAGAACCGACCCAGGGGTCAAAAGTCAACAAGCGGGTGTACCC	6404						
Db	1693	CGTCAAGCCGCTGTCCAGAACCGACCCAGGGGTCAAAAGTCAACAAGCGGGTGTACCC	1752						
QY	6405	TGGCCCAAGCATCATGAGAGTCAATCAACAAGACTACCCCGGACCAACCAAGCAGAC	6464						
Db	1753	TGGCCCAAGCATCATGAGAGTCAATCAACAAGACTACCCCGGACCAACCAAGCAGAC	1812						
QY	6465	TGAGGCAACCCCTGCGCGCCCTCTTACTCTTCCCTGGGGCAGCTGCCCCCTCTGG	6524						
Db	1813	TGAGGCAACCCCTGCGCGCCCTCTTACTCTTCCCTGGGGCAGCTGCCCCCTCTGG	1872						
QY	6525	ACCTCGCGGCCCAACCAAGTGAACCTTACTCCGCGCCCGGACCAAGGTGCCCCGGCCC	6584						
Db	1873	ACCTCGCGGCCCAACCAAGTGAACCTTACTCCGCGCCCGGACCAAGGTGCCCCGGCCC	1932						
QY	6585	GTGCTCCCCCAACGAGGAGGGGCAAGAGTCTTCAGAGCCAAACAAGACGTGCT	6644						
Db	1933	GTGCTCCCCCAACGAGGAGGGGCAAGAGTCTTCAGAGCCAAACAAGACGTGCT	1992						
QY	6645	TGGTGTGTGTGAGAGCGGATTTGAACCTGTGTCCCAACCGGAGGGCATGACGAGCCAG	6704						
Db	1993	TGGTGTGTGTGAGAGCGGATTTGAACCTGTGTCCCAACCGGAGGGCATGACGAGCCAG	2052						
QY	6705	GGCACCTCCCGAGATCTGTGTAACCGCTGTGTACCGGGATGGGAACAAGAGGCCCA	6764						
Db	2053	GGCACCTCCCGAGATCTGTGTAACCGCTGTGTACCGGGATGGGAACAAGAGGCCCA	2112						
QY	6765	GCAGATGGGCTCCAAGTCTCCAGGCAACAACGAGCGCGCAGCCTTCTTCAGCAAC	6824						
Db	2113	GCAGATGGGCTCCAAGTCTCCAGGCAACAACGAGCGCGCAGCCTTCTTCAGCAAC	2172						
QY	6825	TGACGAGAGCAACTCCGCACTGTCAAGTCCAAGAGCAAGATCAACAAGAGCTGA	6884						
Db	2173	TGACGAGAGCAACTCCGCACTGTCAAGTCCAAGAGCAAGATCAACAAGAGCTGA	2232						
QY	6885	ACAACCAACAACGGAGATGAGCTGAATTAACAATAGCCAGCTGGAGCGAGATTCTTA	6944						
Db	2233	ACAACCAACAACGGAGATGAGCTGAATTAACAATAGCCAGCTGGAGCGAGATTCTTA	2292						

OY		6945	ATATGCCCCGCATCATCCGGAAACAGGCTTTATGACTTATGAAGAACCGAGGCGGTGACGAGAC	7004
Db		2293	ATTATGCCCGCATCACCGBMAAGGCTTTATGACTTATGAAGAACCGAGGCGGTGACGAGAC	2355
OY		7005	ATGCGACGACCAACATG9GGGCTGAGGACCATTAATTAGAAGGCACTCATGGGTAAATATG	7064
Db		2353	ATGCGACGACCAACATG9GGGCTGAGGACCATTAATTAGAAGGCACTCATG-----	2402
OY		7065	ACCAGTGGAAGAAGTCCCCGCCGCTCAGCGCCAATGCTTTAACCCTCTGAATGCACTG	7122
Db		2403	-----	2402
OY		7125	CCAGCTGCCCCGTGCTATGCCCATTAACCGCTGTGACGSAAGAGTAGACAACACTCA	7189
Db		2403	-----	2402
OY		7185	CCTGCGCAGGTGGCGGGCGGAAAAGCCAAAGTCTCTGGAGAACCCAGACCGGAAAAAGCA	7244
Db		2403	-----GGTGGCGGCGGAAAGGCCAAAGTCTCTGGAGAACCCAGACCGGAAAAAGCA	2455
OY		7245	AGTCCCGGCCCCCGGGCCCTGGCATCTG9GGAGCCGGACCCCTGTGTCTTCATAGTCACT	7304
Db		2455	AGTCCCGGCCCCCGGGCCCTGGCATCTG9GGAGCCGGACCCCTGTGTCTTCATAGTCACT	2514
OY		7305	CGAGAGGAGACTGCAACCGCGSAGACGCGCTCAACCAACCGCGT9GGAGAGACAGGCGCT	7366
Db		2515	CGAGAGGAGACTGCAACCGCGSAGACGCGCTCAACCAACCGCGT9GGAGAGAGAGGCGCTT	2577
OY		7365	CGTCCGACGTTCCAACGCACTTCCCCTTACAAACCCCTGATCATGCGGCTGACAGCGGGTG	7424
Db		2575	CGTCCGACGTTCCAACGCACTTCCCCTTACAAACCCCTGATCATGCGGCTGACAGGCGGGTG	2634
OY		7425	TCAATGGCTTCCCAACCCCAACCGGGCTTCCCGGGGAGAGGGGCGCTTCCGTGGACCCC	7488
Db		2635	TCAATGGCTTCCCAACCCCAACCGGGCTTCCCGGGGAGAGGGGCGCTTCCGTGGGCGCC	2694
OY		7485	ACCAAGCCTTGGGAGAGAGAGGCCAAAGCACTGCTCTGCTCGACAGTACAGACAACCTCCG	7544
Db		2695	ACCAAGCCTTGGGAGAGAGAGGCCAAAGCACTGCTCTGCTCGACAGTACAGACAACCTCCG	2755
OY		7545	ACACGAGTGACTCAAAACAGGGCGGGGGGGGGCGGCGGTBUERSPERLUTGACAGTCC	7604
Db		2755	ACACGAGTGACTCAAAACAGGGCGGGGGGGGGCGGCGGTBUERSPERLUTGACAGTCC	2806
OY		7605	CAGGAGCGCAACAGAAACGCGCCCTGACAGAGAGGGGGCGGCTCCCATCCCCCAACAAAG	7664
Db		2807	CAGGAGCGCAACAGAAACGCGCCCTGACAGAGAGGGGGCGGCTCCCATCCCCCAACAAAG	2866
OY		7665	AAGGAGCCCCCTGAATCCGCTGCGCTGCGCTCATCTCTGTCGTCAGAGCGGCAATCCTT	7724
Db		2867	AAGGAGCCCCCTGAATCCGCTGCGCTGCGCTCATCTCTGTCGTCAGAGCGGCAATCCTT	2926
OY		7725	GCCCT 7728 	
Db		2927	GCCCT 2930	
<b>RESULT_13</b>				
BC047524				
LOCUS				
DEFINITION Mus musculus nuclear receptor co-repressor 2, mRNA (cDNA clone				
IMAG:4526899), partial cds.				
ACCESSION BC047524				
VERSION BC047524.1 GI:29126781				
KEYWORDS				
SOURCE				
ORGANISM Mus musculus (house mouse)				
MUS musculus				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;				
Sciurognathi; Muridae; Muridae; Murinae; Mus.				
1 (bases 1 to 2964)				
REFERENCE				
AUTHORS Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,				

TITLE	Klausner R.D., Collins F.S., Wagner L., Schenker C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Sherman C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scieczniak T.E., Brownstein M.J., Uddin T.B., Toshitsuki S., Canciani P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abrahamson R.D., Mullany S.J., Boeck S.A., McKean P.J., McKernan K.U., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S., Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E., Jones S.J. and Marra M.A.
JOURNAL	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
REFERENCE	12477932
REMARK	2 (bases 1 to 2964)
COMMENT	Direct Submission Submitted (03-MAR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a> Contact: MGC help desk Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a> Tissue Procurement: The Cepko Laboratory cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: The I.M.A.G.E. Consortium (LNL) Center: Stanford University School of Medicine, Stanford, CA 94305 Web site: <a href="http://www.sngc.stanford.edu">http://www.sngc.stanford.edu</a> Contact: (Dickson, Mark) <a href="mailto:mcdpaxil@stanford.edu">mcdpaxil@stanford.edu</a> Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Series: IRAC Plate: 44 Row: 7 Column: 22 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis
SOURCE	Location/Qualifiers 1. .2964 /organism="Mus musculus" /mol_type="mRNA" /db_xref="taxon:10090" /clone="IMAGE:4526899" /tissue_type="Eye, retina, mouse strain C57BL/6" /clone_id="NIH_MGC_94" /lab_host="DH10B" /note="Vector: pCMV-SPORT6"
gene	<1. .2964 /gene="Ncor2" /note="synonyms: SMRT, SMRTe, N-COR" /db_xref="GeneID:20602" /db_xref="MGI:1337080" <1. .1896 /gene="Ncor2" /codon_start=1 /product="Ncor2 protein" /protein_id="AAH47524.1" /db_xref="GI:29126782" /db_xref="GeneID:20602" /db_xref="MGI:1337080" /translation="VTSVERGPTVLRMARSTSTSPVRAATPPAPKPLIGTLLGVPYTLNRPVILPKETSRVAPRPDADAGAFPTKTPAPSPASSGSSSPRLSPASSTSTLARTPAKNIAPHHASPDPAFTSADLREKTKSPFSIQELNELSLGTHSGA
CDS	

## ORIGIN

GYSPDVEIPISVSPSSLTHDKLSKPLEBELEKSHLBELRHKQBPENKLSABANAHP  
HLRPLBESQPSSSLLOTPQIKQHVVTSAQHSQHSI SEVITDQYTHNHQOQLSGPLAP  
LVSFPGACPVLDLRPPSDYLPPDHDTGTPHSPHSGRSPSPKSTVLSGSSD  
AIEPVSPEGMTEBGRHASTAYPLILYRQBOBPMGSKSPQNTSOPAPSPKLTESN  
SAMYKSKQEIINKLNTNRNRPENYNQPGREIIPMBAITGAGMTCRSAQVQDHS  
TNMGJEAIIIRALMKGIYQWMBPPPLVANAFLPPLNASLPLPAAPITTAQDSQHAL  
TSFGGGGAKVSGRPSRKASPAAGLASGDHPPSVSVSGDNNRRTPLTNRWED  
RPSAGSTPFPYVNPILMRLOAGVMASSPPPGIAMSGLAGHMHAMDEBPXPLCSQY  
ETLSDSB"

Query Match 19.9%; Score 1703.4; DB 9; Length 2964;  
Best Local Similarity 79.1%; Pred. No. 5.8e-212;  
Matches 2347; Conservative 1; Mismatches 511; Indels 108; Gaps 23;

QY 5661 TCACCGCTGTGGAGGCCAGCAAGCCACGAGCTG-----AGGTCACCTCCACT 5711  
DB 2 TCACCTCGGTGAAACCCGACGCGCCACGCTCTGAAGTGGGCGAGGTCCACTCCACT 61  
QY 5712 CCTCACCCGTTGCGCCAGCTGACATTCCACCTGACCAACCACTGCCACTGGCGGCA 5771  
DB 62 CTTGCGCTGTGCGCCAGCTGCAATTCCACCTGACCAACCACTGCCACTGGTGGCA 121  
QY 5772 CCCTGATGGGCTTAACCTTAACCTTAATGAGGCCGCTCTTGTCTGCCCAAGAGGCCCCC 5831  
DB 122 CCTTGAAGGGGCTTAACCTTAACCTTAATGAGGCCGCTCTTGTTCACCAAGAGACCTCTC 181  
QY 5832 GGGTGGCCCGGACAGAGGGGCGCCGAGAGACAGACGGGCGCATGCTTCCCTGGCAAGCCC 5891  
DB 182 GGGTGGCCCGGACAGAGGGGCGCCGAGAGACAGACGGGCGCATGCTTCCCTGGCAAGCCC 241  
QY 5892 CAGCCGCTCCGGGCTGAGAGCCGCTCTCTCCCGAGCAAGAGGCTCGAGAGCCCGCGCC 5951  
DB 242 CGGGCG-----GGAGCCCGCTCTCTCAACCAAGAGAGCTCCGAGCCCGCATGCC 292  
QY 5952 TAGTGCCTCTGTCTGTGCGCACGCGCACCATGCGCGCACCCCTTGGAAGAACTCGCAC 6011  
DB 293 TAGCACCCCGGACGTTCAGGCGCACAGACCATGCGCGCACCCCAAGAAAGAACTTGGCAC 352  
QY 6012 CTACACAGCGCGCGGAGCCGCGCGGCGCGACCTGCGCTCGGCTCGAGCCCGGACCGGG 6071  
DB 353 CCACCAATGCGAGTCCGAGCCCGCGCGG-----CCACCTCGGCTCGAGTTCGACCGAG 409  
QY 6072 AAAAGACTCAAGTAACCTTTTCCATCAGAGAACTGGAACTCCGTTCTCTGGGTAAAC 6131  
DB 410 AAAAGACTCAAGTAACCTTTTCCATCAGAGAACTGGAACTCCGTTCTCTGGGTAAAC 469  
QY 6132 AC---GCGACGCTACAGCCCGGAGGGGTGAGCCGCTCAGCCCTGTGAGCTCAACCA 6188  
DB 470 ACAGTGGAGCTGTGCTACAGCCCGGAGTGGGTGAGGCGCATCAGCCCGGTGAGCTCCCA 529  
QY 6189 GTCTGACCCAGCAAGAGGGGCTCCCAAGACCTGGAAAGAGCTCGACAAAGACCACTGG 6248  
DB 530 GCTTACCCAGCAAGAGGGGCTCTCAACCTCTGGAAGAGCTAAGAAAGACCACTTGG 589  
QY 6249 AGGGGAGCTGCGGCGCAGGAGCGAGCGCCCGTGAAGCTTGGCGGGAGGCGCGCCACC 6308  
DB 590 AAGGGAGCTGCGGCGCAGAGCGAGCGCCCATGAAGCTCAGCGGGAGGCTGGCCCAATC 649  
QY 6309 TCCCACTCTGCGGCGCAGCTGCTGAGAGCGACGCTCTGTCCAGCCCGCTGCTTCAGACCG 6368  
DB 650 TCCCACTCTGCGGCGCAGCTGCTGAGAGCGACGCTCTGTAGCCCACTCTTCAGACG 709  
QY 6369 CCGGAGGGGTCAAAAGTCAAGAGGGGTGTGACCTGCGCCAGACATCAGTGAAGTCA 6428  
DB 710 CCGGAGGGGTCAAAAGTCAAGAGGGGTGTGACCTGCGCTAGACATCAGTGAAGTCA 769  
QY 6429 TCACACAGGACTACACCGGCGACCAACCAAGACGCTCAGCGCACCCCTGCGCGCGCCC 6488  
DB 770 TTAACGAGGACTACACGCGCGCACCAACCGGACGAGTCAAGTGGCCCTTTCGCGCGCCC 829  
QY 6489 TCTACTCTTCCCTTGGGCGCAGCTGCGCGTCTGTGACCTTCGCGCGCCACCAAGTACC 6548

DB 830 TCTACTCTTTCCCGGAGCGAGCTGCGCTGTCTTGAGATCTTTCGCGGCCACCGAGTACC 889  
QY 6549 TCTACTCTCCGCGCCCGAGCATGTGTCCCGGCGCGTGTGCTCCCGCAGCGAAGGGG 6608  
DB 890 TCTACTCTCCACCCCGGACCATGGCACCCGAGCCCGGGATCCCGCAGGTAAAGGGG 949  
QY 6609 GCAAGAGTCTTCAGAGCGCAACCAAGACGTGTGTTGGGTGTGTGTAAGACGTTATTG 6668  
DB 950 GCAAAAGGTCCCGAGAACCCAGCAAAACATGCTGTCTGTGGGAGCGACGACGAGATGCATTG 1009  
QY 6669 AACCTGTCTCCCGACCGAGGGGCAATGAGGAGCGAGGCACTCCCGAGTGTGTATCC 6728  
DB 1010 AGCTGTGTCCCGACAGAGGCGATGACTGAGCCAGGACATGTCTGAGACACTGTATCC 1069  
QY 6729 CGTGTCTTACCGGAGTGGGAAACAGACGAGCCGAGAGATGGGCTTCAAGTCTCAG 6788  
DB 1070 CACTGTCTTATGAGAGCGGGGAAACAGGGCGAGCC---CAGGATGGGCTTAAAGTCTCAG 1126  
QY 6789 GCAACACAGCGCGAGCGCGCATGCTTTCAGCAGAGTGAACGAGGCAATTCGCGCATGG 6848  
DB 1127 GCAACACAGCGCGAGCGCGCATGCTTTCAGTAAAGCTGATGAGAGCAATTCGCGCATGG 1186  
QY 6849 TCAAGTCCAGAGGAGAGATCAACAGAAAGCTGAACACCCACACCGGATGAGCTTG 6908  
DB 1187 TGAAGTGAAGAGAGGAGAGATCAACAGAAATCTCAACACCAACCGGAGAGCGAG 1246  
QY 6909 AATACATATGACGCGAGCTGTGAGACGAGATCTTCAATATGATGCGCGCATCAACGGAACAG 6968  
DB 1247 AATACATATGATGCGAGCTGTGAGACGAGATCTTCAATATGATGCGCGCATCAACGGAACAG 1306  
QY 6969 GCCTTATGACCTATGAGACCGAGCGGTGTGAGAACATGCGACGACCAATGTGGGCTGG 7028  
DB 1307 GCCTTATGACCTGTGAGAGCGAGCGGTGTGAGAACAGCGGACGACCAATGTGGGCTGG 1366  
QY 7029 AGGCGATATTAAGAAAGGACATCATGGGTAAATATGACAGAGGGGAGAGTCCCGCGCG 7088  
DB 1367 AGGCGATATTAAGAAAGGACATCATGGGTAAATATGACAGAGGGGAGAGTCCCGCGCG 1426  
QY 7089 TCAGCGCAATGCTTTTAAACCTCTGAATGCGCAGTGCAGCTGTGCC---GCTGTATATG 7145  
DB 1427 TCAGCGCAATGCTTTTAAACCTCTGAATGCGCAGTGCAGCTGTGCC---GCTGTATATG 1486  
QY 7146 CCATTAACGCTGTGACGAGCGAGAGTGAACACACACTCACTTCGCGAGGTGCGCGGGA 7205  
DB 1487 CCATTAACGCTGTGACGAGCGAGAGTGAACACACACTCACTTCGCGAGGTGCGCGGGA 1546  
QY 7206 AGGCGAAGTCTGTGGGAGAGCCGAGCGAGCCGAAAGCCAAAGTCCCGCGCGCGGCTGG 7265  
DB 1547 AAGCGAAGTCTGTGGGAGAGCCGAGCGAGCCGAAAGCCAAAGTCCCGAGCACGAGCTAG 1606  
QY 7266 CATGTGGGAGCGGACCACTCTGTCTCTCTGATGCACTCGAGGGAGACTGCAACGCGC 7325  
DB 1607 CGTCCGAGAGCGAGCCCTTCTGTCTCTCTGATGCACTCGAGGGGAGCTGCAATCGCC 1666  
QY 7326 GAGCGCGCTCAACACGCGGTGTGGAGAGACAGGCGCTCTGTCCGAGTTTCAGGCAAT 7385  
DB 1667 GAAACACACTCAACCAACGCTGTGTGGAGAGCGCGCGCTCTGATGCAAGGCTCAAGCAT 1726  
QY 7386 TCCCTTAACACCCCGTAACTAGCGGCTGAGGGGGGTGATGAGCTTCCCAACCCCAAC 7445  
DB 1727 TCCCTTAACACCCCGTAACTAGCGGCTGAGGGGGGTGATGAGCTTCCCAACCCCAAC 1786  
QY 7446 CGGCGCTCCCGCGGAGCGAGCGGCGCTCTGTGAGCGCCCAACGCGCTTGGAGCGAGAGC 7505  
DB 1787 CTGGCTTGGCGGAGGAGCGAGCGGCGCTTGTAGCTGTCTCCACCAAGCGCTTGGAGTGAAGGC 1846  
QY 7506 CCAAGCGACTGTCTGTCTGCGAGTACGAGACCTTTCGACAGGAGTGAATCAAGACAG 7565  
DB 1847 CCAAGCGACTGTGTGTTCAGAGTATGAGACACTTCGAGACGAGGAGTGAACAGGATTTG 1906  
QY 7566 GCGGGGGGGGGCGGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 7625  
DB 1907 GCGGGGGGGGGCGGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1948

QY 7626 CTGACAGAGCGGGGCGGCTGCGACTCCCCCAACCAAGGAAGAGCCCTGAGTCCGCT 7685  
 DB 1949 CAGCATGAGACACACAGCTCTACTATCCCAAGACTGAGAGAGAGCCCTGAGTCTGCT 2008  
 QY 7686 GCGCCTCCATC-----CATCTGTCCGTCCAGAGCGGCATCTCTGCTG 7729  
 DB 2009 GCGCGTCAATCCGTCCGCTCCGTCCACTCATCTGTCCATCCAGAGCTGGCATCC-TGCGTG 2067  
 QY 7730 TCTAAAGCCTTAATAAGATCTCCCGCGCGGCTGCGCTCTGTCAG----ACCTTACTCA 7785  
 DB 2068 TCTAAAGCCTTAATAAGATCTCCCGCGCGGCTGCGCTCTGCGCATGACCTTAACACTCA 2127  
 QY 7786 GGGGATGTTTAACTGAGTCTGGGAAAGGAGGGGAAAGGGGCGG----- 7829  
 DB 2128 GGGGATGTTTAACTGAGTCTGAGAGGGGAGTGAACAGAGGAGGAGGAGACAGCGGCGCAG 2187  
 QY 7830 --GAGAGGGGACACGACAGCGGTGTGGACCCACACACAGGCGGACAGGCGGCGCAGGAC 7887  
 DB 2188 GAGGGGGGGGACAGCAATCTGTGTGTCAGTCCGACTGTGCATCTGGGCGAG--CGGGGAC 2246  
 QY 7888 CCAAGCAGGATACCAACGCACTCCAGCGCACTGCT--CCCCGAATGATTTGGAAC 7945  
 DB 2247 CCAACACAGGCTGACAGGACCTCCATGACCGCTCGCTCCCTTAAACGCAATTTGGAAC 2306  
 QY 7946 CAAAGCTAACTGAGTCTGCGACGCCCGCGGCTCCGCTCCGCTCCATCCGCTTACG 8005  
 DB 2307 CAAAGCTAACTGAGTCTGCGAGTGTCTGCGCTCCGCTCCGCTCCGCTCCGCTTACG 2366  
 QY 8006 GCTCTGACAGATGACACGACGCGCTGTCCAGCGCCCGACGCTCGTCCGCTCCGCTCCAC 8065  
 DB 2367 GCTCTGACAGACAGACGTTTCCAGCTTATCCGCTTATGCTGTGATCATGACAGTCT 2426  
 QY 8066 AGACTGCCCCAGCAACAGATGCTGGAACCAAGTCAGGCCAGAGTGGCGGACCAAAAG 8125  
 DB 2427 CCAAAAGGCCACCCAGGCCACACAGACTGGAGCCCA-TCAACACAGAGTGGTGCAC-AAGG 2484  
 QY 8126 GGCAGAGTGGGCTGGGGGGAAACGATGCTCCGAGACGACGCTGTTTTTTCACACAT 8185  
 DB 2485 GGCAGAGCTGGGCT--GGGACACGAGTGTTCAGAGACCTGACCGTTCGCGCTGT 2542  
 QY 8186 CGTTGCGCAGCGGTGGGAAGAAAGCAGATGTAATGATGTGTGTTTACAGGGTAT 8245  
 DB 2543 TGCCTGTGGCAACGG-----AGGAAAGCAGTGTAAATG--GTGTGGCTTACAGGGTAT 2596  
 QY 8246 ATTTTGTATCTTCAATGAATTAATTCAGATGTTTACGCAAGAAAGACTTACCAGT 8305  
 DB 2597 ATTTTGTATCTTCAATGAATTAATTCAGATGTTTACGCAAGAAAGACTTACCAGT 8365  
 QY 8306 ATTAAGTCTGCTGTGCTTTTGTATCTGCTTACCGTTCAGAGCGGTGACGAGCGAC 8365  
 DB 2656 GTTTCGCGCTGTGCTGCTTGTCTTACCTTACCTGTTTTCAGAGCGGTGACGAGCGC-AA 2712  
 QY 8366 GTGCGTGAACCCCATCACTGACAGACCAAGGGGGGCGGAGCTGCTGTCAGCGCCGCTG 8425  
 DB 2713 GGGGCTGGCCCAACATAGCAGAGACTTGGGGGTTCAGGGGCTCCGACACGCGACCTGTG 2772  
 QY 8426 -TGCTCTCCCTCCCTCCCTTCTGTTGGCAGAAATGAATTCGATTTCTGTGGCGCC 8484  
 DB 2773 CCGTTCGCCCCCAACCGCTTCTTGGGCAAAATGAACTTATGATCGTATTCGTGGCGCT 2832  
 QY 8485 ATTTGGGAGGAGGTGTGTATTTCTGTCACTTACACAGCTGCTTCAATTTAAAGGAAT 8544  
 DB 2833 CTCTGTGCA--CGGGCGCATTTCTGTCACTTACACAGTGTTCATTTAAGAAAGCAAT 2889  
 QY 8545 TATACTCCAAAAAATTTTAAAAAATTTTAAAAA 8571  
 DB 2890 ATTAAGTGAAGACACGTTAAAAAATTTTAAAAA 2916

RESULT 14  
 AC073916/c AC073916 205283 bp DNA linear PRI 27-MAR-2003

DEFINITION Homo sapiens 12 BAC RP11-408118 (Roswell) Park Cancer Institute Human BAC Library) complete sequence.  
 ACCESSION AC073916  
 VERSION AC073916.41 GI:29293398  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 205283)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T., Barbarella,J., Benton,J., Bimase,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhey,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carion,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Barnhart,C., Edgar,D., Edwards,C.C., Elhai,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,K., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Hollway,C., Hollins,B., Homsi,P., Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshihara,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Ieal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Iicharge,O., Ieal,C., Liu,J., Liu,W., Louieged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Mathewar,M., Mapa,P., Marondei,I., Martin,R., Martindale,A., Martinez,B., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Mescher,S., Metzker,K., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabac,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwoun,G., Ogunyemi,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shoohtari,N., Sison,I., Sodergren,B., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,R., Thomas,S., Uemami,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Kucheriapali,R., Weinstock,G., and Gibbs,R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 205283)  
 Worley,K.C.  
 Direct Submission  
 Submitted (05-JUL-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 205283)  
 Worley,K.C.  
 Direct Submission  
 Submitted (31-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 4 (bases 1 to 205283)  
 Worley,K.C.  
 Direct Submission  
 Submitted (01-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA







3913	5368:	contig of 1456 bp in length
5369	5468:	gap of 100 bp
5469	6519:	contig of 1051 bp in length
6520	6619:	gap of 100 bp
6620	7852:	contig of 1233 bp in length
7853	7952:	gap of 100 bp
7954	8464:	contig of 512 bp in length
8465	8564:	gap of 100 bp
8565	10347:	contig of 1783 bp in length
10348	10447:	gap of 100 bp
10448	12145:	contig of 1698 bp in length
12146	12245:	gap of 100 bp
12246	14006:	contig of 1761 bp in length
14007	14106:	gap of 100 bp
14107	15358:	contig of 1252 bp in length
15359	15458:	gap of 100 bp
15459	17152:	contig of 1694 bp in length
17163	17252:	gap of 100 bp
17253	18800:	contig of 1548 bp in length
18801	18900:	gap of 100 bp
18901	20291:	contig of 1391 bp in length
20292	20391:	gap of 100 bp
20392	21889:	contig of 1498 bp in length
21890	21989:	gap of 100 bp
21990	24355:	contig of 2366 bp in length
24356	24455:	gap of 100 bp
24456	26757:	contig of 2302 bp in length
26758	26857:	gap of 100 bp
26858	28317:	contig of 1460 bp in length
28318	28417:	gap of 100 bp
28418	30781:	contig of 2364 bp in length
30782	30881:	gap of 100 bp
30882	32789:	contig of 1908 bp in length
32790	32889:	gap of 100 bp
32890	35214:	contig of 2325 bp in length
35215	35314:	gap of 100 bp
35315	37181:	contig of 1867 bp in length
37182	37281:	gap of 100 bp
37282	39642:	contig of 2361 bp in length
39643	39742:	gap of 100 bp
39743	41978:	contig of 2236 bp in length
41979	42078:	gap of 100 bp
42079	44663:	contig of 2885 bp in length
44664	45063:	gap of 100 bp
45064	47265:	contig of 2202 bp in length
47266	47365:	gap of 100 bp
47366	49947:	contig of 2482 bp in length
49948	49947:	gap of 100 bp
49948	52574:	contig of 2527 bp in length
52475	52574:	gap of 100 bp
52575	54767:	contig of 2193 bp in length
54768	54867:	gap of 100 bp
54868	57938:	contig of 3071 bp in length
57939	58038:	gap of 100 bp
58039	60538:	contig of 2500 bp in length
60540	60638:	gap of 100 bp
60639	63322:	contig of 2784 bp in length
63423	63322:	gap of 100 bp
63523	66038:	contig of 2516 bp in length
66039	66138:	gap of 100 bp
66140	69185:	contig of 3047 bp in length
69186	69285:	gap of 100 bp
69286	72321:	contig of 3036 bp in length
72322	72421:	gap of 100 bp
72422	75240:	contig of 2819 bp in length
75241	75340:	gap of 100 bp
75341	77598:	contig of 2258 bp in length
77599	77698:	gap of 100 bp
77699	82485:	contig of 4787 bp in length
82486	82585:	gap of 100 bp
82586	86177:	contig of 3592 bp in length
86178	86278:	gap of 100 bp
86278	91129:	contig of 4852 bp in length

	*	91130	91239:	gap of 100 bp	in length
	*	91230	94475:	contig of 3246 bp	in length
	*	94476	94575:	gap of 100 bp	in length
	*	94576	98382:	contig of 3807 bp	in length
	*	98383	98482:	gap of 100 bp	in length
	*	98483	103891:	contig of 5409 bp	in length
	*	103892	103991:	gap of 100 bp	in length
	*	103992	109749:	contig of 5758 bp	in length
	*	109750	109849:	gap of 100 bp	in length
	*	109850	116075:	contig of 6226 bp	in length
	*	116076	116175:	gap of 100 bp	in length
	*	116176	122437:	contig of 6262 bp	in length
	*	122438	122537:	gap of 100 bp	in length
	*	122538	132382:	contig of 9855 bp	in length
	*	132393	132492:	gap of 100 bp	in length
	*	132493	141881:	contig of 9389 bp	in length
	*	141882	141981:	gap of 100 bp	in length
	*	141982	161970:	contig of 19989 bp	in length.
FEATURES					
source					
			1..161970		
		/organism="Homo sapiens"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:9606"			
		/chromosome="12"			
		/map="12"			
		/clone="RP11-665C13"			
		/clone_lib="RPC1-11 Human Male BAC"			
		1..1052			
misc_feature		/note="assembly_fragment"			
		1053..1152			
gap		/estimated_length=100			
		1153..2378			
misc_feature		/note="assembly_fragment"			
		2379..2478			
gap		/estimated_length=100			
		2479..3812			
misc_feature		/note="assembly_fragment"			
		3813..3912			
gap		/estimated_length=100			
		3913..5368			
misc_feature		/note="assembly_fragment"			
		5369..5468			
gap		/estimated_length=100			
		5469..6519			
misc_feature		/note="assembly_fragment"			
		6520..6619			
gap		/estimated_length=100			
		6620..7852			
misc_feature		/note="assembly_fragment"			
Query Match		11.1%;	Score 948.4;	DB 14;	Length 161970;
Best Local Similarity		97.6%;	Pred. No. 3.7e-114;		
Matches 965;		Conservative 4;	Mismatches 19;	Indels 1;	Gaps 1;
Dy	7574	GCGGCGGGCGGTGTTTCTCAAGTCCTCCAGGAGGCACAGGAAGGCCCTTGACAGA	7633		
	:: ::				
Db	60639	GCGGCGGGCGGGGCGGCGGTGTCAAGTCCTCCAGGAGGCACAGGAAGGCCCTTGACAGA	606398		
Dy	7634	GCGGCGGCGGTGCGCATCCCCAACAAAGGAAGAGCCCGTAGTCGCGGCGCTCC	7693		
Db	60699	GCAAGGCGGTGCGCATCCCCAACAAAGGAAGAGCCCGTAGTCGCGGCGCTCC	60758		
Dy	7694	ATCATCTGTGCGTCAGAGCGGCAATCTTCTGTCTTAAGCCTTAATAAGACTCCC	7753		
Db	60759	ATCATCTGTGCGTCAGAGCGGCAATCTTCTGTCTTAAGCCTTAATAAGACTCCC	60818		
Dy	7754	GCCCCGAGCTGCGCTGTGCAACCTTAATCAAGGAGATGTTTACTGTGCTCGGAAAG	7813		
Db	60819	GCCCCGAGCTGCGCTGTGCAACCTTAATCAAGGAGATGTTTACTGTGCTCGGAAAG	60878		
Dy	7814	GAGGGGAAGGCGCGGGGAGGGGACCGCACGCTGTGTGGCAGCCACAACAGGCGGCCA	7873		
Db	60879	GAGGGGAAGGCGCGGGGAGGGGACCGCACGCTGTGTGGCAGCCACAACAGGCGGCCA	60938		



**THIS PAGE BLANK (USPTO)**